GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 13:52:37 ; Search time 96 Seconds
(without alignments)
1560.378 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 3817
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1495.5	39.2	744	5	Q9V8I8	Q9v8i8 drosophila
2	1337.5	35.0	634	11	Q9D687	Q9d687 mus musculu
3	1326	34.7	615	11	O88576	O88576 mus musculu
4	1323	34.7	616	4	Q8TF10	Q8tf10 homo sapien
5	1307.5	34.3	592	4	Q9NP91	Q9np91 homo sapien
6	1305	34.2	662	5	Q9W4S0	Q9w4s0 drosophila
7	1303.5	34.1	592	11	Q8VDB9	Q8vdb9 mus musculu
8	1302.5	34.1	616	11	Q64093	Q64093 rattus norv
9	1302.5	34.1	635	11	Q91WT6	Q91wt6 mus musculu
10	1300.5	34.1	635	11	O88575	O88575 mus musculu
11	1294.5	33.9	615	11	Q62687	Q62687 rattus norv
12	1282	33.6	628	4	Q96N87	Q96n87 homo sapien
13	1141	29.9	577	11	Q91XG6	Q91xg6 mus musculu
14	1032	27.0	791	11	Q91ZQ2	Q91zq2 mus musculu
15	1024	26.8	397	4	Q9BYZ7	Q9byz7 homo sapien
16	1008.5	26.4	631	5	Q9NB97	Q9nb97 drosophila
17	1008.5	26.4	631	5	Q961H9	Q961h9 drosophila
18	1007	26.4	629	13	Q9QZV1	Q9Qzv1 brachydanio
19	1001	26.2	576	5	Q9V7R0	Q9v7r0 drosophila
20	1000.5	26.2	642	4	Q9UN76	Q9un76 homo sapien
21	996.5	26.1	633	11	Q8VC47	Q8vc47 mus musculu
22	987.5	25.9	638	11	Q9D317	Q9d317 mus musculu
23	987.5	25.9	638	11	Q9JMA9	Q9jma9 mus musculu
24	987.5	25.9	638	11	Q91Y60	Q91y60 mus musculu
25	986.5	25.8	617	6	Q9MYW8	Q9myw8 macaca mula
26	985.5	25.8	628	4	Q96KH8	Q96kh8 homo sapien
27	983	25.8	630	13	O42482	O42482 rana catesb
28	976.5	25.6	619	11	Q9JJ41	Q9jj41 mus musculu
29	976.5	25.6	619	11	Q9R0X6	Q9r0x6 mus musculu
30	975.5	25.6	617	11	Q8R2I2	Q8r2i2 mus musculu
31	974.5	25.5	602	4	Q8TCC2	Q8tcc2 homo sapien
32	969.5	25.4	635	13	Q9DGN5	Q9dgn5 gallus gall
33	965.5	25.3	1201	5	Q9W1J0	Q9w1j0 drosophila
34	965	25.3	617	11	Q63380	Q63380 rattus norv
35	964.5	25.3	638	11	Q9R183	Q9r183 mus musculu
36	961	25.2	289	4	Q9H9F5	Q9h9f5 homo sapien
37	961	25.2	289	4	Q8TBM6	Q8tbm6 homo sapien
38	961	25.2	640	11	Q8VBW1	Q8vbw1 mus musculu
39	958	25.1	620	6	Q9GJT6	Q9gjt6 macaca fasc
40	955.5	25.0	515	13	Q9I8Q2	Q9i8q2 coturnix co
41	954	25.0	620	6	Q9GJT5	Q9gjt5 saimiri sci
42	953	25.0	611	13	Q91502	Q91502 torpedo mar
43	952.5	25.0	567	11	Q9WTR3	Q9wtr3 rattus norv
44	952.5	25.0	597	11	Q9WTR4	Q9wtr4 rattus norv
45	952.5	25.0	614	11	Q8VCS9	Q8vcs9 mus musculu

ALIGNMENTS

RESULT 1
Q9V8I8
ID Q9V8I8 PRELIMINARY; PRT; 744 AA.
AC Q9V8I8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG5226 protein.
GN CG5226.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003799; AAF57678.1; -.
DR   FlyBase; FBgn0034355; CG5226.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   744 AA;  82953 MW;  ED4A742B3E0E483B CRC64;

  Query Match             39.2%;  Score 1495.5;  DB 5;  Length 744;
  Best Local Similarity   44.0%;  Pred. No. 1.5e-112;
  Matches  279;  Conservative  124;  Mismatches  198;  Indels   33;  Gaps    6;

Qy      59 DRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLEL 118
           :| ||: |:|: |: ||:|||||||||||||||:|||||:|:|:::|:||: ||||| :||
Db      54 ERAAWSGKMQFFLSIIGYSVGLGNIWRFPYLCQQNGGGAFLIPFMVMLILEGIPLFLIEL 113

Qy     119 AVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPW 178
           :|||:| |::|||: | | ||||| ||||| ||| ||||||| |  || | ||:|||||
Db     114 GIGQRMRLGALGVWNTIHPWLGGIGISSCIVTLFVALYYNVIITWVFFYLFNSFRYPLPW 173

Qy     179 SECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIG 238
           | ||:  ||:   :| || ||| |||||||  || : |: | ||| | : |||:: |:|
Db     174 SSCPL--NGTGFELE-ECAKSSETTYFWYRTTLDAAPSMDEPGGLKWWIVLCLMLSWTIV 230

Qy     239 GMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVW 298
           | |||||||||:||:|||||:||  | :||: |||| |::||:|||: |:|:| ||
Db     231 FFIVMKGIQSSGKVVYFTSLFPYIVLTIFFIRGITLRGAGAGLMHMYTPKVEKLLEPTVW 290

Qy     299 REVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFK 358
           : |||||:  || || :| | |||   |||  |  |||  |  |:: |::|:| :||||
Db     291 LDAATQVFYSFGLAFGSLIAFGSYNTPKNNCVRDVLLVSVCNAVTAIYASVVIFAILGFK 350

Qy     359 ANIMNEKCVVENAE-----KILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVK 413
           | :  ::||  | |     |:|| ||                    |    :  :  :
Db     351 ATVNVDRCVASNTEILVKNKLLG-LN--------------------DTQGYEQAMSLLN 388

Qy     414 EDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGS 473
           :  | | | |  ||| : :||||||| ||:|:   | :|||:|:|| ||::|||||
Db     389 GTELLRLQLSECSLAHELDNAAEGTGLAFIVFTQAIVELPGAPFWAVLFFTMLLSLGLGS 448

Qy     474 MIGTMAGITTPIID---TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATL 530
           || : |:   : |     :| |:  |   |:| |:|| :|   :| |:: ||| :: |:
Db     449 QIGILEGMLCTLFDIDIIKRVKKQHVTGVVCLFCFIVGFIFCTGAGEYWLKMFDSFAGTI 508

Qy     531 PLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQ 590
           | :: ::| ||| :||| ::| ::: :| |:|| |:: : |::: |: |  :  :|::
Db     509 GLVVVALMEMIAVIFIYGHERFTEDIFQMTGYRPGRYWQWTWRYIGPVIMVCILVSSVVF 568

Qy     591 LGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNT 650
           : :  | | ||  :         :||| | : :::|:   ||:|:||::| |  |     :
Db     569 MVIRNPTYGAWNADLGMIEQKSYPNWVMGIALSMILAGVLPMPIVFLMRSFQCLKVDLDI 628

Qy     651 LSVSYKKARMMKDISNLEENDETRFILSKVPSEA 684
             | ::         : :||:        | ::
Db     629 HQGSIRRNETTASTKEMIDNDDDNMSPDMPPQDS 662


RESULT 2
Q9D687
ID   Q9D687      PRELIMINARY;      PRT;   634 AA.
AC   Q9D687;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   4632401C08Rik protein.
GN   4632401C08RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=SKIN;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK014544; BAB29422.1; -.
DR   MGD; MGI:1921588; 4632401C08Rik.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   634 AA;  71366 MW;  433EB23C97E9FE5B CRC64;

  Query Match             35.0%;  Score 1337.5;  DB 11;  Length 634;
  Best Local Similarity   42.3%;  Pred. No. 8e-100;
  Matches  268;  Conservative  117;  Mismatches  213;  Indels   35;  Gaps   12;

Qy      49 KAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLII 108
           : :|:| :|  || |::| ||:|  :|| |||||:|||||||| :||||:::|:|:||:
Db      22 EVIEKE-EAGSRPKWDNKAQYMLTCVGFCVGLGNVWRFPYLCQSHGGGAFMIPFLILLVF 80

Qy     109 IGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYF 168
           |||| :|| |:|||:|:||:|||  | | | ||| :|  |  ||||| || | ::||
Db      81 EGIPLLYLEFAIGQRLRKGSMGVWSSIHPALKGIGIASMFVSFMVGLYYNTIIAWVMWYF 140

Qy     169 FKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMT 228
           | ||| ||||||||: :|     || || |||:  ||||||  |:|| |||:|| : | :
Db     141 FNSFQEPLPWSECPLNQN-QTGYVE-ECAKSSSVDYFWYRETLNISTSISDSGSIQWWIL 198

Qy     229 LCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPK 288
           |||   ||:  : : :||:::|| :| :|  |||||  ||:||| |:|| :||: :|||
Db     199 LCLTCAWSVLYVCIIRGIETTGKAVYITSTLPYVVLTIFLIRGLTLKGATNGIVFLFTPN 258

Qy     289 LVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLAT 348
           : :: :|  | :   |||:  | |||:| |||||   |||  |  :|| || |||| |
Db     259 ITELSNPNTWLDAGAQVFYSFSLAFGGLISFSSYNSVHNNCEMDSVIVSVINGFTSVYAA 318
```

```
Qy    349 LVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSR----DLIPPHVNFSHLTTKDYME 404
          ||: ::||:|    : ||           |||:|:     ||  |  | :    : | :
Db    319 TVVYSIIGFRATERFDDCV----------NTNILTLINGFDL--PEGNVTSENFEAYQQ 365

Qy    405 MDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFL 464
          |       ::  |    | :   | : |:|||||| ||||:|  | || |||:||:
Db    366 WCN---ATNPQAYAQLKFQTCDINSFLSEGVEGTGLAFIVFTEAITKMPVSPLWSVLFFI 422

Qy    465 MLINLGLGSMIGTMAGITTPIIDTS----KVPKEMFTVGCCVFTFLVGLLFVQRSGNYFV 520
          ||  ||| || | | |:  |: | :    | |||: |   |: |:|:  :|   || |::
Db    423 MLFCLGLSSMFGNMEGVVVPLQDLNITPKKWPKELLTGLICLGTYLIAFIFTLNSGQYWL 482

Qy    521 TMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCM 580
          :: | :: ::|| :|   |  || ::||  :| :::  |:| :|  |:   |: |||| |
Db    483 SLLDSFAGSIPLLIIAFCEMFAVVYVYGVDRFNKDIEFMIGHKPNIFWQVTWRVVSPLIM 542

Qy    581 AVL-TTASIIQLGVTPPAYSAW---IKEEAAERYLYFPNWPMALLITLIVVATLPIPVVF 636
           |:     :|::  |   || |    :|    : : :|||  |:::|:  |  | ||
Db    543 LVIFLFFFVIEVNKT-LMYSIWDPNYEEFPKSQKIPYPNWVYAVVVTVAGVPCLSIPCFA 601

Qy    637 V---LRHFHLLSDGSNTLSVSYKKARMMKDISN 666
          :   :|:    ||  : |  :   | :  |: |
Db    602 IYKFIRNCCQKSDDHHGLVNTLSTASVNGDLKN 634


RESULT 3
O88576
ID   O88576      PRELIMINARY;      PRT;   615 AA.
AC   O88576; O88577; O88578; O88579; O88580; O88581;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium- and chloride-dependent transporter XTRP2.
GN   XTRP2 OR XT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC   TISSUE=KIDNEY;
RX   MEDLINE=99131080; PubMed=9932288;
RA   Nash S.R., Giros B., Kingsmore S.F., Kim K.M., El-Mestikawy S.,
RA   Dong Q., Fumagalli F., Seldin M.F., Caron M.G.;
RT   "Cloning, gene structure, and genomic localization of an orphan
RT   transporter from mouse kidney with six alternatively-spliced
RT   isoforms.";
RL   Recept. Channels 6:113-128(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: 6 ISOFORMS; A12 (SHOWN HERE), A11, B11, A10,
CC       B9 AND A8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
DR   EMBL; AF075262; AAC27757.1; -.
DR   EMBL; AF075263; AAC27758.1; -.
DR   EMBL; AF075264; AAC27759.1; -.
DR   EMBL; AF075265; AAC27760.1; -.
DR   EMBL; AF075266; AAC27761.1; -.
DR   EMBL; AF075267; AAC27762.1; -.
DR   MGD; MGI:1336892; Xtrp2.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; FALSE_NEG.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Alternative splicing.
FT   DOMAIN       1     26       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     27    47       1 (POTENTIAL).
FT   DOMAIN       48    52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53    73       2 (POTENTIAL).
FT   DOMAIN       74    83       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     84    104      3 (POTENTIAL).
FT   DOMAIN       105   105      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     106   126      4 (POTENTIAL).
FT   DOMAIN       127   177      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     178   198      5 (POTENTIAL).
FT   DOMAIN       199   206      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     207   227      6 (POTENTIAL).
FT   DOMAIN       228   288      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     289   309      7 (POTENTIAL).
FT   DOMAIN       310   397      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     398   418      8 (POTENTIAL).
FT   DOMAIN       419   441      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     442   462      9 (POTENTIAL).
FT   DOMAIN       463   472      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     473   493      10 (POTENTIAL).
FT   DOMAIN       494   520      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     521   541      11 (POTENTIAL).
FT   DOMAIN       542   570      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     571   591      12 (POTENTIAL).
FT   DOMAIN       592   615      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     353   353      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     377   422      SASGPGLAFIVFTEAVLHMPGASVWSVLFFGMLFTLGLSSM
FT                               FGNME -> VLLCGLCSSLGCCLPWVCPPCLGTWRVSLHHY
FT                               WTWGSYPKVYPRRS (IN ISOFORM A8).
FT   VARSPLIC     377   444      SASGPGLAFIVFTEAVLHMPGASVWSVLFFGMLFTLGLSSM
FT                               FGNMEGVITPLLDMGILPKGIPKEVMT -> PTWKQISGAR
FT                               VLGEGCARLTSRVCEASVLP (IN ISOFORM B11 AND
FT                               ISOFORM B9).
FT   VARSPLIC     445   454      MISSING (IN ISOFORM A11).
FT   VARSPLIC     499   552      FCDDIEWMTGRRPGLYWQVTWRVVSPMLLFGIFLSYIVLLI
FT                               QTPPSYKAWNPQY -> NIFPQERRSSTQAGCRSPVCSCPS
FT                               CPHCGSLELLWLSYCPSTNRGGRLRIWKVV (IN
FT                               ISOFORM A10 AND ISOFORM B9).
FT   VARSPLIC     553   615      MISSING (IN ISOFORM A10).
SQ   SEQUENCE   615 AA;  69228 MW;  A7925C0397FC63F8 CRC64;

  Query Match           34.7%;  Score 1326;  DB 11;  Length 615;
  Best Local Similarity 42.0%;  Pred. No. 6.6e-99;
  Matches  271;  Conservative  121;  Mismatches  200;  Indels   54;  Gaps   14;

Qy     43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
          :| |    |: |   ::|| |::||||:|: |||:|||||||||||||| :||||:|:||
Db      3 QASGMDPLVDIE---DERPKWDNKLQYLLSCIGFAVGLGNIWRFPYLCQTHGGGAFLIPY 59

Qy    103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
           : |:  |||||::|||:|||:|||||||  | | |||:|     |   | |||| ::
Db     60 FIALVFEGIPLFYIELAIGQRLRRGSIGVWKTISPYLGGVGLGCFSVSFLVSLYYNTVLL 119

Qy    163 WSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGG 222
          | :::|  |||:|||| ||:  | :  |   ||: |   :|||||: |:|:  || :|
Db    120 WVLWFFLNSFQHPLPWSTCPLDLNRTGFV--QECQSSGTVSYFWYRQTLNITSDISNTGT 177

Qy    223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
          : ||: |||:  ||   : | :||:|:|||:||::||||:||  ||:||| | || :|::
Db    178 IQWKLFLCLVACWSTVYLCVIRGIESTGKVIYFTALFPYLVLTIFLIRGLTLPGATEGLI 237

Qy    283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
          ::||| :  : :|:|| : |||:|| | | ||| | |:|||   |||  |  ::: :|
Db    238 YLFTPNMKTLQNPRVWLDAATQIFFSLSLAFGGHIAFASYNPPRNNCEKDAVIIALVNSM 297

Qy    343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDY 402
          ||: |:: :| |:||||:    :|:  |   ||  :|     ||  | :: |    :|
Db    298 TSLYASIAIFSVMGFKASNDYGRCLDRN---ILSLIN----EFDL--PELSIS----RD- 343

Qy    403 MEMDNVIMTVKEDQ---FSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWS 459
          |  :|:|  :   |    : | |  | ||| ||||  | |||| |||: | | :  ||
Db    344 -EYPSVLMYLNATQTARVAQLPLKTCHLEDFLDKSASGPGLAFIVFTEAVLHMPGASVWS 402

Qy    460 VMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRS 515
          |:|| ||  ||| || | | |: ||::|       :|||: |   |  ||  : |  :|
Db    403 VLFFGMLFTLGLSSMFGNMEGVITPLLDMGILPKGIPKEVMTGVICFACFLSAICFTLQS 462
```

```
Qy    516 GNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFV 575
Db    463 GGYWLEIFDSFAASLNLIIFAFMEVVGVIHIYGMKRFCDDIEWMTGRRPGLYWQVTWRVV 522

Qy    576 SPLCMAVLTTASIIQLGVTPPAYSAW--------IKEEAAERYLYFPNWPMALLITLIVV 627
Db    523 SPMLLFGIFLSYIVLLIQTPPSYKAWNPQYEHFPSREEK-----FYPGWVQVTCVLLSFL 577

Qy    628 ATLPIPVVFVLRHFHLLSDGSNTLSVSYK---KARMMKDISNLEEN 670
Db    578 PSLWVPGVALAQ---LLS--------QYKQRWKATHLESGLKLQES 612


RESULT 4
Q8TF10
ID   Q8TF10      PRELIMINARY;      PRT;   616 AA.
AC   Q8TF10;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Neurotransmitter transporter RB21A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou Y., Yu L., Zhao S.Y.;
RT   "Cloning of a new human cDNA similar to Rattus norvegicus
RT   neurotransmitter transporter rB21a.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ding J.B., Yu L., Zhao S.Y.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF125107; AAL75944.1; -.
SQ   SEQUENCE   616 AA;  68365 MW;  D3DF0C282A1D0237 CRC64;

  Query Match             34.7%;  Score 1323;  DB 4;  Length 616;
  Best Local Similarity   41.7%;  Pred. No. 1.2e-98;
  Matches  263;  Conservative  125;  Mismatches  199;  Indels   44;  Gaps   11;

Qy     25 LALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIW 84
Db      3 LAIKKPASCDPRAGAERAEAGAMEKA---------RPLWANSLQFVFACISYAVGLGNVW 53

Qy     85 RFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGF 144
Db     54 RFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGV 113

Qy    145 SSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTY 204
Db    114 ASVVVSFFLSMYYNVINAWAFWYLFHSFQDPLPWSVCPL--NGNHTGYDEECEKASSTQY 171

Qy    205 FWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVL 264
Db    172 FWYRKTLNISPSLQENGGVQWEPALCLLLAWLVVYLCILRGTESTGKVVYFTASLPYCVL 231

Qy    265 ACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNK 324
Db    232 IIYLIRGLTLHGATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNE 291

Qy    325 QDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLS 384
Db    292 PSNNCQKHAIIVSLINSFTSIFASIVTFSIYGFKATFNYENCL-----KKVSLLLTNTFD 346

Qy    385 RDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAF 442
Db    347 -------LEDGFLTASNLEQVKGYLASAYPSKYSEMFPQIKNCSLESELDTAVQGTGLAF 399

Qy    443 IAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTV 498
```

```
Db    400 IVYTEAIKNMEVSQLWSVLYFFMLLMLGIGSMLGNTAAILTPLTDSKIISSHLPKEAISG 459

Qy    499 GCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTE 558
Db    460 LVCLVNCAIGMVFTMEAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLKA 519

Qy    559 MLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQLGVTPPAYSAWIKEEAAERYLY--- 612
Db    520 MTGRAVSWYWKVMWAGVSPLLIVSLFVFYLSDYILTGTL--KYQAW---DASQGQLVTKD 574

Qy    613 FPNWPMALLITLIVVATLPIPV----VFVLR 639
Db    575 YPAYALAVIGLLVASSTMCIPLAALGTFVQR 605


RESULT 5
Q9NP91
ID   Q9NP91      PRELIMINARY;      PRT;   592 AA.
AC   Q9NP91; Q9NQ77; Q9NPQ2; O75590;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium- and chloride-dependent transporter XTRP3.
GN   XTRP3 OR XT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RA   Kiss H.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-315 FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=99131080; PubMed=9932288;
RA   Nash S.R., Giros B., Kingsmore S.F., Kim K.M., El-Mestikawy S.,
RA   Dong Q., Fumagalli F., Seldin M.F., Caron M.G.;
RT   "Cloning, gene structure, and genomic localization of an orphan
RT   transporter from mouse kidney with six alternatively-spliced
RT   isoforms.";
RL   Recept. Channels 6:113-128(1998).
RN   [3]
RP   SEQUENCE OF 232-592 FROM N.A.
RA   Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; XT3 (SHOWN HERE) AND XT3A; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: KIDNEY AND SMALL INTESTINE.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
DR   EMBL; AJ276207; CAB99310.1; -.
DR   EMBL; AJ276208; CAB99311.1; -.
DR   EMBL; AJ289880; CAB96872.1; -.
DR   EMBL; AL389979; CAB97535.1; -.
DR   EMBL; AF075260; AAC27755.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; FALSE_NEG.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Alternative splicing.
FT   DOMAIN        1     42       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     43     63       1 (POTENTIAL).
FT   DOMAIN       64     79       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     80    100       2 (POTENTIAL).
FT   DOMAIN      101    107       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    108    128       3 (POTENTIAL).
FT   DOMAIN      129    165       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    166    186       4 (POTENTIAL).
```

```
FT   DOMAIN       187     194       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     195     215       5 (POTENTIAL).
FT   DOMAIN       216     241       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     242     262       6 (POTENTIAL).
FT   DOMAIN       263     276       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     277     297       7 (POTENTIAL).
FT   DOMAIN       298     389       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     390     410       8 (POTENTIAL).
FT   DOMAIN       411     431       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     432     452       9 (POTENTIAL).
FT   DOMAIN       453     465       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     466     486       10 (POTENTIAL).
FT   DOMAIN       487     504       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     505     525       11 (POTENTIAL).
FT   DOMAIN       526     554       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     555     575       12 (POTENTIAL).
FT   DOMAIN       576     592       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     131     131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     357     357       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     195     231       MISSING (IN ISOFORM XT3A).
FT   CONFLICT      31      31       F -> N (IN REF. 2).
FT   CONFLICT     217     217       L -> V (IN REF. 2).
SQ   SEQUENCE   592 AA;  65913 MW;  AF6E94E19895377F CRC64;

  Query Match             34.3%;  Score 1307.5;  DB 4;  Length 592;
  Best Local Similarity   42.8%;  Pred. No. 2e-97;
  Matches  255;  Conservative  121;  Mismatches  185;  Indels   35;  Gaps   10;

Qy      60 RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELA 119
           || | : ||:: | | ::||||||:|||||||| |||::|||||:::||: |:|| :|||
Db       5 RPLWANSLQFVFACISYAVGLGNVWRFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELA 64

Qy     120 VGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWS 179
           ||||:|:|||| |  | | | |:| :| :|  |: :|||||  |: :| | ||| |||||
Db      65 VGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFLSMYYNVINAWAFWYLFHSFQDPLPWS 124

Qy     180 ECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGG 239
            ||:  ||:    : ||||:|:| |||||: |:|| |: |:||: |:  ||||: | :
Db     125 VCPL--NGNHTGYDEECEKASSTQYFWYRKTLNISPSLQENGGVQWEPALCLLLAWLVVY 182

Qy     240 MAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWR 299
           : : :| :|:|||:||::  || ||  :|:||| | || :|:::|||||: :: :|: |
Db     183 LCILRGTESTGKVVYFTASLPYCVLIIYLIRGLTLHGATNGLMYMFTPKIEQLANPKAWI 242

Qy     300 EVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKA 359
            |||:|| |||||| :| |:|||:  |||     :|| || |||: |::| | : ||||
Db     243 NAATQIFFSLGLGFGSLIAFASYNEPSNNCQKHAIIVSLINSFTSIFASIVTFSIYGFKA 302

Qy     360 NIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSA 419
              | |:     | :  | ||          :    ||  :  ::   : :    ::|
Db     303 TFNYENCL-----KKVSLLLTNTFD-------LEDGFLTASNLEQVKGYLASAYPSKYSE 350

Qy     420 L--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
           :   :  | || ||| :||||||||| :|||: :   |  |||::| ||: ||:||||:|
Db     351 MFPQIKNCSLESELDTAVQGTGLAFIVYTEAIKNMEVSQLWSVLYFFMLLMLGIGSMLGN 410

Qy     478 MAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLT 533
           | | ||: |    :| :|||  :   |:    :|::|   :|||:  :|:||:||| |
Db     411 TAAILTPLTDSKIISSHLPKEAISGLVCLVNCAIGMVFTMEAGNYWFDIFNDYAATLSLL 470

Qy     534 LIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQ 590
           |||::| ||| ::|| ::|  :|  | |     ::  ||  |||| :    |   :  |
Db     471 LIVLVETIAVCYVYGLRRFESDLKAMTGRAVSWYWKVMWAGVSPLLIVSLFVFYLSDYIL 530

Qy     591 LGVTPPAYSAWIKEEAAERYLY---FPNWPMALLITLIVVATLPIPV----VFVLR 639
           |     | ||   :|::  |    :| : :|::  |:  :|: ||:     || |
Db     531 TGTL--KYQAW---DASQGQLVTKDYPAYALAVIGLLVASSTMCIPLAALGTFVQR 581


RESULT 6
Q9W4S0
ID   Q9W4S0      PRELIMINARY;      PRT;   662 AA.
AC   Q9W4S0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CG10804 protein.
GN   CG10804.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003427; AAF45873.2; -.
DR   FlyBase; FBgn0029663; CG10804.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   662 AA;  73629 MW;  58C3A9AFF042FB86 CRC64;

  Query Match             34.2%;  Score 1305;  DB 5;  Length 662;
  Best Local Similarity   40.7%;  Pred. No. 3.7e-97;
  Matches  248;  Conservative  109;  Mismatches  209;  Indels   44;  Gaps    4;

Qy      34 KQSVLNVAGEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
           |  :: :|  :  :|  || | : |:| :|:||: ::|| ||::|||||:|||||| |||
Db      28 KAELVAIATASADQQDIVEGEKEGEERESWDSKIMFLLATIGYAVGLGNVWRFPYLAQKN 87

Qy      94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
           ||||:|||| ::| | |||:|:||||:|||:|:|:|:|||  : | |||| || :|  |
Db      88 GGGAFLVPYFIMLCIQGIPIFYLELAIGQRLRKGAIGVWSQVSPYLGGIGISSAVVSYIV 147
```

```
Qy    154 GLYYNVIIGWSIFYFFKSFQYPLPWSECPV-VRNGSVAVVEAECEKSSATTYFWYREALD 212
Db    148 ALYYNTIIAWCLIYLLHSFESPLPWADCPTRLYKNFTYDHEPECVASSPTQFYWYRTTLQ 207

Qy    213 ISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGL 272
Db    208 CSESVDMPENFNYHMAIALIVSWFLVYICMVQGITSSGKIVYMTAIFPYVVLIIFFFRGI 267

Qy    273 LLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFD 332
Db    268 TLKGAADGVAHLFTPRWETLLDPVVWLEAGTQIFFSLGLAFGGLIAFSSYNPANNNCYRD 327

Qy    333 GALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHV 392
Db    328 AILVSLTNCGTSMFAGVVVFSVIGFKATATFDRCTEERNGLVA----------------Q 371

Qy    393 NFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHF 452
Db    372 NKTH------------------------NLPVCDLQTELANSASGTGLAFIIFTEAINQF 407

Qy    453 PTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTS---KVPKEMFTVGCCVFTFLVGL 509
Db    408 PGAQLWAVLFFLMLFTLGIDSQFGTLEGVVTSLVDMKLFPNLPKEYIVGALCFSCCTISM 467

Qy    510 LFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYF 569
Db    468 CFANGAGSYIFQLMDSFAGNFPLLIIALFECLSISYIYGVRRFSDDIEMMTGSRPNFYWM 527

Qy    570 YMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVAT 629
Db    528 FCWKYLSPCAMVTILLASFYQLLTEGSSYPAWIGSKGATEGMEWPHWCIVVAFFLILSSI 587

Qy    630 LPIPVVFVLR 639
Db    588 LWIPIVAVLR 597
```

RESULT 7
Q8VDB9
ID   Q8VDB9   PRELIMINARY;   PRT;   592 AA.
AC   Q8VDB9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   X transporter protein 3 similar 1.
GN   XTRP3S1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RA   Kiss H., Darai E., Kiss C., Kost-Alimova M., Klein G., Dumanski J.P.,
RA   Imreh S.;
RT   "Comparative Sequence Analysis of the Common Eliminated Region 1
RT   (C3CER1) on Human Chromosome 3p21.3 with the Orthologous Regions on
RT   Mouse Chromosome 9F.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ428067; CAD20989.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   592 AA; 66178 MW; 547B9F8C0B7A99AA CRC64;

  Query Match   34.1%; Score 1303.5; DB 11; Length 592;
  Best Local Similarity   44.0%; Pred. No. 4.2e-97;
  Matches 259; Conservative 115; Mismatches 183; Indels   31; Gaps   11;

```
Qy     60 RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELA 119
Db      5 RPQWGHPLQFVFACISYAVGLGNVWRFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELA 64

Qy    120 VGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWS 179
Db     65 VGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFLSMYYNVINAWGFWYLFHSFQDPLPWS 124

Qy    180 ECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGG 239
Db    125 VCPL--NSNHTGYDEECEKASSTQYFWYRKTLNISPSIQENGGVQWEPALCLTLAWLMVY 182

Qy    240 MAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWR 299
Db    183 LCILRGTESTGKVVYFTASMPYCVLIIYLVRGLTLHGATNGLMYMFTPKMEQLANPKAWI 242

Qy    300 EVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKA 359
Db    243 NAATQIFFSLGLGFGSLIAFASYNEPSNNCQKHAIIVSIINSSTSIFASIVTFSIYGFKA 302

Qy    360 NIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSA 419
Db    303 TFNYENCL----NKVI-LLLTN--SFDL-----EDGFLTVSNLEEVKNYLASTYPNKYSE 350

Qy    420 L--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
Db    351 VFPHIRNCSLESELDTAVQGTGLAFIVYTEAIKNMEVSQLWSVLYFFMLLMLGIGSMLGN 410

Qy    478 MAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLT 533
Db    411 TAAILTPLTDSKVISSYLPKEAISGLVCLINCAVGMVFTMEAGNYWFDIFNDYAATLSLL 470

Qy    534 LIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQ 590
Db    471 LIVLVETIAVCYVYGLKRFESDLRAMTGRTLSWYWKVMWAFVSPLLIVGLFIFYLSDYIL 530

Qy    591 LGVTPPAYSAWIKEEAAERYLYFPNWP---MALLITLIVVATLPIPVV 635
Db    531 TGTL--QYQAW---DATQGQLVTKDYPPHALAVIGLLVASSTMCIPLV 573
```

RESULT 8
Q64093
ID   Q64093   PRELIMINARY;   PRT;   616 AA.
AC   Q64093;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium- and chloride-dependent transporter XTRP3 (NEUROTRANSMITTER
DE   transporter RB21A).
GN   XTRP3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE=95094940; PubMed=8001687;
RA   Smith K.E., Fried S.G., Durkin M.M., Gustafson E.L., Borden L.A.,
RA   Branchek T.A., Weinshank R.L.;
RT   "Molecular cloning of an orphan transporter. A new member of the
RT   neurotransmitter transporter family.";
RL   FEBS Lett. 357:86-92(1995).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC   FAMILY (SNF).
DR   EMBL; S76742; AAB32806.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; FALSE_NEG.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.

```
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     37       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     38     58       1 (POTENTIAL).
FT   DOMAIN       59     66       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     67     87       2 (POTENTIAL).
FT   DOMAIN       88    103       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    104    124       3 (POTENTIAL).
FT   DOMAIN      125    189       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    190    210       4 (POTENTIAL).
FT   DOMAIN      211    218       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    219    239       5 (POTENTIAL).
FT   DOMAIN      240    265       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    266    286       6 (POTENTIAL).
FT   DOMAIN      287    300       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    301    321       7 (POTENTIAL).
FT   DOMAIN      322    413       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    414    434       8 (POTENTIAL).
FT   DOMAIN      435    455       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    456    476       9 (POTENTIAL).
FT   DOMAIN      477    489       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    490    510       10 (POTENTIAL).
FT   DOMAIN      511    533       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    534    554       11 (POTENTIAL).
FT   DOMAIN      555    578       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    579    599       12 (POTENTIAL).
FT   DOMAIN      600    616       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    155    155       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    381    381       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   616 AA;  68971 MW;  6B0DAB04315CE829 CRC64;

  Query Match             34.1%;  Score 1302.5;  DB 11;  Length 616;
  Best Local Similarity   43.1%;  Pred. No. 5.3e-97;
  Matches  265;  Conservative  122;  Mismatches  191;  Indels   37;  Gaps   14;

Qy      38 LNVAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
Db       3 LAIKRRASRGQRPGPDEKRARDMEKARPQWGNPLQFVFACISYAVGLGNVWRFPYLCQMY 62

Qy      94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
Db      63 GGGSFLVPYLIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFL 122

Qy     154 GLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDI 213
Db     123 SMYYNVINAWGFWYLFHSFQDPLPWSVCPL--NSNRTGYDEECEKASSTQYFWYRKTLNI 180

Qy     214 SDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLL 273
Db     181 SPSIQENGGVQWEPALCLTLAWLMVYLCILRGTESTGKVVYFTALMPYCVLIIYLVRGLT 240

Qy     274 LRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDG 333
Db     241 LHGATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNEPSNDCQKHA 300

Qy     334 ALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVN 393
Db     301 VIVSVINSSTSIFASIVTFSIYGFKATFNYENCL----NKVI-LLLTN--SFDL-----E 348

Qy     394 FSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTH 451
Db     349 DGFLTASNLEEVKDYLASTYPNKYSEVFPHIRNCSLESELNTAVQGTGLAFIVYAEAIKN 408

Qy     452 FPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLV 507
Db     409 MEVSQLWSVLYFFMLLMLGMGSMLGNTAAILTPLTDSKVISSYLPKEAISGLVCLINCAV 468

Qy     508 GLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRF 567
Db     469 GMVFTMEAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLRAMTG-RPLNW 527

Qy     568 YF-YMWKFVSPLC---MAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWP---MAL 620
```

```
Db     528 YWKAMWAFVSPLLIIGLFIFYLSDYILTGTL--QYQAW---DATQGQLVTKDYPPHALAV 582

Qy     621 LITLIVVATLPIPVV 635
Db     583 IGLLVASSTMCIPLV 597


RESULT 9
Q91WT6
ID   Q91WT6      PRELIMINARY;      PRT;   635 AA.
AC   Q91WT6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   X transporter protein 3.
GN   XTRP3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013484; AAH13484.1; -.
DR   MGD; MGI:1336891; Xtrp3.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   635 AA;  70660 MW;  EC913793A807EC7B CRC64;

  Query Match             34.1%;  Score 1302.5;  DB 11;  Length 635;
  Best Local Similarity   41.3%;  Pred. No. 5.5e-97;
  Matches  266;  Conservative  132;  Mismatches  197;  Indels   49;  Gaps   15;

Qy      26 ALEEPVDYKQSVLNVAGEAG---GKQKAVE-------EELDAEDRPAWNSKLQYILAQIG 75
Db       8 AVSLPEDEELQPWGGAGGPGQHPGRPRSTECAHPGVVEKV----RPKWDNPLQFLLVCIS 63

Qy      76 FSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYI 135
Db      64 YAVGLGNVWRFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTI 123

Qy     136 CPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAE 195
Db     124 SPYLSGVGIASLVVSFLASVYFNVINTWALWYLFHSFQDPLPWSVCPL--NSNHTGYDEE 181

Qy     196 CEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYF 255
Db     182 CEKASSTQYFWYRKTLNISPSIQENGGVQWEPALCLTLAWLMVYLCILRGTESTGKVVYF 241

Qy     256 SSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGG 315
Db     242 TTSLPYFVLIIYLVRGLTLHGATNGLAYMFTPKIEQLANPKAWINAATQIFFSLGLGCGG 301

Qy     316 VIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKIL 375
Db     302 LIAFASYNEPSNDCQKHALIVSVINSTTAIFSSIVTFSIYGFKATFNYENCL----NKVI 357

Qy     376 GYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDK 433
Db     358 -LLLTN--SFDL-----EDGFLTVSNLEEVKNYLASTYPNKYSEVFPHIRNCSLESELDT 409

Qy     434 SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TS 489
Db     410 AVQGTGLAFIVYTEAIKNMEVSQLWSVLYFFMLLTLGMGSMVGTGTAILTPLTDSKIISS 469

Qy     490 KVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGP 549
Db     470 YLPKEAISGLVCLLNCAIGMVFTMEAGNYWFDLFNDYTATLSLLLIVLVETIAVCYVYGL 529
```

```
Qy     550 KKFMQELTEMLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQLGVTPPAYSAWIKEEA 606
           |:|  :|  | |      ::  || ||||| :    :  :  |  |     | ||   :|
Db     530 KRFESDLRAMTGRTLSWYWKVMWAFVSPLLIVGLFIFYLSDYILTGTL--QYQAW---DA 584

Qy     607 AERYLY---FPNWPMALLITLIVVATLPIPVV----FVLRHFHL 643
             : ::    :| : :|::  |:  :|: ||:|    || ||| :
Db     585 TQGHVVTKDYPTYALAVIGLLVASSTMCIPLVALGTFVTRHFKI 628


RESULT 10
O88575
ID   O88575      PRELIMINARY;      PRT;   635 AA.
AC   O88575;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium- and chloride-dependent transporter XTRP3.
GN   XTRP3 OR XT3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=KIDNEY;
RX   MEDLINE=99131080; PubMed=9932288;
RA   Nash S.R., Giros B., Kingsmore S.F., Kim K.M., El-Mestikawy S.,
RA   Dong Q., Fumagalli F., Seldin M.F., Caron M.G.;
RT   "Cloning, gene structure, and genomic localization of an orphan
RT   transporter from mouse kidney with six alternatively-spliced
RT   isoforms.";
RL   Recept. Channels 6:113-128(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- TISSUE SPECIFICITY: KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
DR   EMBL; AF075261; AAC27756.1; -.
DR   MGD; MGI:1336891; Xtrp3.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; FALSE_NEG.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     56       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      57    77       1 (POTENTIAL).
FT   DOMAIN        78    85       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      86    106      2 (POTENTIAL).
FT   DOMAIN        107   127      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      128   148      3 (POTENTIAL).
FT   DOMAIN        149   208      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      209   229      4 (POTENTIAL).
FT   DOMAIN        230   237      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      238   258      5 (POTENTIAL).
FT   DOMAIN        259   284      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      285   305      6 (POTENTIAL).
FT   DOMAIN        306   319      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      320   340      7 (POTENTIAL).
FT   DOMAIN        341   432      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      433   453      8 (POTENTIAL).
FT   DOMAIN        454   474      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      475   495      9 (POTENTIAL).
FT   DOMAIN        496   508      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      509   529      10 (POTENTIAL).
FT   DOMAIN        530   547      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      548   568      11 (POTENTIAL).
FT   DOMAIN        569   597      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      598   618      12 (POTENTIAL).
FT   DOMAIN        619   635      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      174   174      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      400   400      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   635 AA;  70641 MW;  EC9272C12507EC7B CRC64;

  Query Match             34.1%;  Score 1300.5;  DB 11;  Length 635;
  Best Local Similarity   41.3%;  Pred. No. 8e-97;
  Matches  266;  Conservative  131;  Mismatches  198;  Indels   49;  Gaps   15;

Qy      26 ALEEPVDYKQSVLNVAGEAG---GKQKAVE-------EELDAEDRPAWNSKLQYILAQIG 75
           |:  | | :       ||  |   |: :: |       |::    || |:: ||::|  |
Db       8 AVSLPEDEELQPWGGAGGPGQHPGRPRSTECAHPGVVEKV----RPKWDNPLQFLLVCIS 63

Qy      76 FSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYI 135
           ::|||||:||||||||  |||::||||:::||: |:|| :|||||||||:| |||| |  |
Db      64 YAVGLGNVWRFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELAVGQRMRXGSIGAWRTI 123

Qy     136 CPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAE 195
             | | |:| :| :|     :|:||| |:::| | ||| ||||| ||:  | :    : |
Db     124 SPYLSGVGIASLVVSFLASVYFNVINTWALWYLFHSFQDPLPWSVCPL--NSNHTGYDEE 181

Qy     196 CEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYF 255
           |||:|:| |||||: |:|| || |:||: |:  ||| : | :  : : :| :|:|||:||
Db     182 CEKASSTQYFWYRKTLNISPSIQENGGVQWEPALCLTLAWLMVYLCILRGTESTGKVVYF 241

Qy     256 SSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGG 315
           ::  || ||  :|||| | || :|: :|||||: :: :|: |   |||:|| |||| ||
Db     242 TTSLPYFVLIIYLVRGLTLHGATNGLAYMFTPKIEQLANPKAWINAATQIFFSLGLGCGG 301

Qy     316 VIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKIL 375
           :| |:|||:  |:|     :|| ||  |:: :::| | : ||||    | |:    |::
Db     302 LIAFASYNEPSNDCQKHALIVSVINSTTAIFSSIVTFSIYGFKATFNYENCL----NKVI 357

Qy     376 GYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDK 433
             | ||  | ||          ||  :  |: | : :   :::| :   :  | || |||
Db     358 -LLLTN--SFDL-----EDGFLTVSNLEEVKNYLASTYPNKYSEVFPHIRNCSLESELDT 409

Qy     434 SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TS 489
           :|||||||||| :|||: :   |  |||::| ||: ||:|||:||   | ||: |    :|
Db     410 AVQGTGLAFIVYTEAIKNMEVSQLWSVLYFFMLLTLGMGSMVGTGTAILTPLTDSKIISS 469

Qy     490 KVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGP 549
           :|||  :   |:   :|::|  :|||:  :|:||:|||  | |||::| ||| ::||
Db     470 YLPKEAISGLVCLLNCAIGMVFTMEAGNYWFDLFNDYTATLSLLLIVLVETIAVCYVYGL 529

Qy     550 KKFMQELTEMLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQLGVTPPAYSAWIKEEA 606
           |:|  :|  | |      ::  || ||||| :    :  :  |  |     | ||   :|
Db     530 KRFESDLRAMTGRTLSWYWKVMWAFVSPLLIVGLFIFYLSDYILTGTL--QYQAW---DA 584

Qy     607 AERYLY---FPNWPMALLITLIVVATLPIPVV----FVLRHFHL 643
             : ::    :| : :|::  |:  :|: ||:|    || ||| :
Db     585 TQGHVVTKDYPTYALAVIGLLVASSTMCIPLVALGTFVTRHFKI 628


RESULT 11
Q62687
ID   Q62687      PRELIMINARY;      PRT;   615 AA.
AC   Q62687;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Renal osmotic stress-induced Na-Cl organic solute cotransporter.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY CORTEX;
RX   MEDLINE=95029937; PubMed=7943364;
RA   Wasserman J.C., Delpire E., Tonidandel W., Kojima R., Gullans S.R.;
RT   "Molecular characterization of ROSIT, a renal osmotic stress-induced
RT   Na(+)-Cl(-)-organic solute cotransporter.";
RL   Am. J. Physiol. 267:F688-F694(1994).
```

```
DR   EMBL; U12973; AAC13771.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   615 AA;  69556 MW;  F41E7DE1DC276918 CRC64;

  Query Match             33.9%;  Score 1294.5;  DB 11;  Length 615;
  Best Local Similarity   42.0%;  Pred. No. 2.4e-96;
  Matches  253;  Conservative  121;  Mismatches  189;  Indels   39;  Gaps   10;

Qy      43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
           :| |    |: |   ::|| |::|||:|: |||:||||||||||||||  :||||:|:|| 
Db       3 QASGMDPLVDIE---DERPKWDNKLQYLLSCIGFAVGLGNIWRFPYLCHTHGGGAFLIPY 59

Qy     103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
            : |:  |||||::|||:|||:||||||||  | | |||:|     |   | |||| |: 
Db      60 FIALVFEGIPLFYIELAIGQRLRRGSIGVWKTISPYLGGVGLGCFSVSFLVSLYYNTILL 119

Qy     163 WSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGG 222
           | :::|  |||:|||| ||:  | :  |   ||: |   :|||||: |:|:  || :| 
Db     120 WVLWFFLNSFQHPLPWSTCPLDLNRTGFV--QECQSSGTVSYFWYRQTLNITSDISNTGT 177

Qy     223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
           : ||: |||:  |:   : | :||:|:|||:||::|||||:||  ||:||| | || :|: 
Db     178 IQWKLFLCLVACWTTVYLCVIRGIESTGKVIYFTALFPYLVLTIFLIRGLTLPGATEGLT 237

Qy     283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
           ::||| :  : : :|| : |||:|| | | ||: | |:|||:  |||  |   :: :| 
Db     238 YLFTPNMKILQNSRVWLDAATQIFFSLSLAFGGHIAFASYNQPRNNCEKDAVTIALVNSM 297

Qy     343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLT-TKD 401
           ||: |:: :| ::||||:    :|:  |   ||  :|              :|  |: ::| 
Db     298 TSLYASITIFSIMGFKASNDYGRCLDRN---ILSLIN-----------EFDFPELSISRD 343

Qy     402 YMEMDNVIM---TVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFW 458
             |  :|:|     : :: : | |  | ||| |||  | |||| |||: | | :  | 
Db     344 --EYPSVLMYLNATQPERVARLPLKTCHLEDFLDKSASGPGLAFIVFTEAVLHMPGASVW 401

Qy     459 SVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQR 514
           ||:|| ||  ||| || | | |: ||: |       |||  |   |   ||  : |  : 
Db     402 SVLFFGMLFTLGLSSMFGNMEGVITPLFDMGILPKGVPKETMTGVVCFICFLSAICFTLQ 461

Qy     515 SGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKF 574
           ||:|:: :|| ::|:| | :   :| : |  :|| |:|  ::  | | ||  ::   |: 
Db     462 SGSYWLEIFDSFAASLNLIIFAFMEVVGVIHVYGIKRFCDDIEWMTGRRPSLYWQVTWRV 521

Qy     575 VSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPN-----WPMALLITLIVVAT 629
           |||: :  :  : |: |  : |:| ||       :| :||:     :|  : :| :::: 
Db     522 VSPMLLFGIFLSYIVLLAQSSPSYKAW-----NPQYEHFPSREEKLYPGWVQVTCVLLSF 576

Qy     630 LP 631
           ||
Db     577 LP 578


RESULT 12
Q96N87
ID   Q96N87      PRELIMINARY;      PRT;   628 AA.
AC   Q96N87;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ31236 fis, clone KIDNE2004828, moderately similar to Mus
DE   musculus orphan transporter isoform A12 (Xtrp2) mRNA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA   Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA   Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA   Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK055798; BAB71018.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   628 AA;  70924 MW;  90EA71CD12C6DC9D CRC64;

  Query Match             33.6%;  Score 1282;  DB 4;  Length 628;
  Best Local Similarity   42.2%;  Pred. No. 2.5e-95;
  Matches  251;  Conservative  111;  Mismatches  189;  Indels   44;  Gaps   10;

Qy      58 EDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
           ::|| |::| ||:|: |||:|||||||||||||  ||||:|:||:: |:  |||:| :| 
Db      16 DERPKWDNKAQYLLSCIGFAVGLGNIWRFPYLCQTYGGGAFLIPYVIALVFEGIPIFHVE 75

Qy     118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
           ||:|||:|:||:|||  | | | |:|     :   : |||| |: | ::|   |||:||| 
Db      76 LAIGQRLRKGSVGVWTAISPYLSGVGLGCVTLSFLISLYYNTIVAWVLWYLLNSFQHPLP 135

Qy     178 WSECP--VVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVW 235
           || ||  : | | |     ||: ||| :||||||: |:|:  |::|| : | : :||  | 
Db     136 WSSCPPDLNRTGFV----EECQGSSAVSYFWYRQTLNITADINDSGSIQWWLLICLAASW 191

Qy     236 SIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDP 295
           ::  | | :||:::|||:||::|||||:||  ||:||| | ||  |::::||| :  : :| 
Db     192 AVVYMCVIRGIETTGKVIYFTALFPYLVLTIFLIRGLTLPGATKGLIYLFTPNMHILQNP 251

Qy     296 QVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVL 355
           :|| : |||:|| | | ||| | |:|||   |:|  |  ::: :|  ||: |:? || || 
Db     252 RVWLDAATQIFFSLSLAFGGHIAFASYNSPRNDCQKDAVVIALVNRMTSLYASIAVFSVL 311

Qy     356 GFKANIMNEKCVVENAEKILGYLN-----TNVLSRDLIPP---HVNFSHLTTKDYMEMDN 407
           ||||    | |:  |   ||  :|        :|||  |    |:|                
Db     312 GFKATNDYEHCLDRN---ILSLINDFDFPEQSISRDDYPAVLMHLN-------------- 354

Qy     408 VIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLI 467
                 : : |  |||| ||||  | ||||: |||   | | :|  |:::|| ||    
Db     355 ---ATWPKRVAQLPLKACLLEDFLDKSASGPGLAFVVFTETDLHMPGAPVWAMLFFGMLF 411

Qy     468 NLGLGSMIGTMAGITTPIIDTSK----VPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMF 523
           |||  :| ||:  : ||::|       ||||  |   |:  ||    |  :||||:: :| 
Db     412 TLGLSTMFGTVEAVITPLLDVGVLPRWVPKEALTGLVCLVCFLSATCFTLQSGNYWLEIF 471

Qy     524 DDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVL 583
           |:::|:| | ::  || : | ::|| |:|  ::  | | ||  ::   |: |||| : : 
Db     472 DNFAASLNLLMLAFLEVVGVVYVYGMKRFCDDIAWMTGRRPSPYWRLTWRVVSPLLLTIF 531

Qy     584 TTASIIQLGVTPPAYSAWIKE----EAAERYLYFPNWPMALLITLIVVATLPIPV 634
             | || |   |  | ||  :     : :  || | |  |  : | ::  | :||
Db     532 -VAYIILLFWKPLRYKAWNPKYELFPSRQEKLY-PGWARAACVLLSLLPVLWVPV 584


RESULT 13
Q91XG6
ID   Q91XG6      PRELIMINARY;      PRT;   577 AA.
AC   Q91XG6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to X transporter protein 2.
GN   XTRP2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010748; AAH10748.1; -.
DR   MGD; MGI:1336892; Xtrp2.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   577 AA;  65254 MW;  5F84F7C5CF1B14E3 CRC64;

  Query Match             29.9%;  Score 1141;  DB 11;  Length 577;
  Best Local Similarity   37.9%;  Pred. No. 6e-84;
  Matches  243;  Conservative  122;  Mismatches  193;  Indels   84;  Gaps   16;

Qy      43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
           :| |    |: |   ::|| |::||||:|: |||:|||||||||||||| :||||:|:||
Db       3 QASGMDPLVDIE---DERPKWDNKLQYLLSCIGFAVGLGNIWRFPYLCQTHGGGAFLIPY 59

Qy     103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
           : |:  |||||::|||:|||:||||||||  | | |||:|     |   | |||| ::
Db      60 FIALVFEGIPLFYIELAIGQRLRRGSIGVWKTISPYLGGVGLGCFSVSFLVSLYYNTVLL 119

Qy     163 WSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGG 222
           | :::|  |||:|||||  ||:  | :  |   ||: |   :|||||: |:|:  || :|
Db     120 WVLWFFLNSFQHPLPWSTCPLDLNRTGFV--QECQSSGTVSYFWYRQTLNITSDISNTGT 177

Qy     223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
           : ||: |||:  ||   : | :||:|:|||:||::||||:||  ||:||| | || :|::
Db     178 IQWKLFLCLVACWSTVYLCVIRGIESTGKVIYFTALFPYLVLTIFLIRGLTLPGATEGLI 237

Qy     283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
           ::||| :  : :|:|| : |||:|| | | ||| | |:|||   |||  |  ::: :|
Db     238 YLFTPNMKTLQNPRVWLDAATQIFFSLSLAFGGHIAFASYNPPRNNCEKDAVIIALVNSM 297

Qy     343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDY 402
           ||: |:: :| |:||||:    :|:  |   ||  :|      ||  | :: |    :|
Db     298 TSLYASIAIFSVMGFKASNDYGRCLDRN---ILSLIN----EFDL--PELSIS----RD- 343

Qy     403 MEMDNVIMTVKEDQ---FSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWS 459
           |  :|:| :   |    : | |  | ||| |||                      | |
Db     344 -EYPSVLMYLNATQTARVAQLPLKTCHLEDFLDK----------------------PTWK 380

Qy     460 VMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYF 519
           :   ::  |       | :|: : : | :|  :     |  ||  : |  :|| |:
Db     381 QISGARVLGEG-------CARLTSRVCEASVLPGVI-----CFACFLSAICFTLQSGGYW 428

Qy     520 VTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLC 579
           : :|| ::|:| | :   :| : |  ||| |:|  ::  | | ||  ::   |: |||:
Db     429 LEIFDSFAASLNLIIFAFMEVVGVIHIYGMKRFCDDIEWMTGRRPGLYWQVTWRVVSPML 488

Qy     580 MAVLTTASIIQLGVTPPAYSAW--------IKEEAAERYLYFPNWPMALLITLIVVATLP 631
           :  :  : |: |  |||:| ||         :||      ::| |     : |  : :|
Db     489 LFGIFLSYIVLLIQTPPSYKAWNPQYEHFPSREEK-----FYPGWVQVTCVLLSFLPSLW 543

Qy     632 IPVVFVLRHFHLLSDGSNTLSVSYK---KARMMKDISNLEEN 670
           :| | : :   |||         ||   ||  ::    |:|:
Db     544 VPGVALAQ---LLS--------QYKQRWKATHLESGLKLQES 574


RESULT 14
Q91ZQ2
ID   Q91ZQ2      PRELIMINARY;      PRT;   791 AA.
AC   Q91ZQ2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glycine transporter type 2.
GN   GLYT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RA   Liu Q.-R., Li Q.-F.;
RT   "Cloning and expression of mouse glycine transporter 2 (Glyt2).";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF411042; AAL17054.1; -.
DR   InterPro; IPR001920; Asp/Glu_race.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00923; ASP_GLU_RACEMASE_1; UNKNOWN_1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; UNKNOWN_1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   791 AA;  87029 MW;  8AA8BB6D0D89EAEF CRC64;

  Query Match             27.0%;  Score 1032;  DB 11;  Length 791;
  Best Local Similarity   32.3%;  Pred. No. 6.2e-75;
  Matches  229;  Conservative  117;  Mismatches  208;  Indels  154;  Gaps   16;

Qy      43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
           | | : ||         |  |:||| :||: :|::|||||:||||||  :|||||:|:||
Db     177 EQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQNGGGAFLIPY 227

Qy     103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
           |::| : |:|:||||:::||   :| : ||  | | | | | :  |: : : :||||||
Db     228 LMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVLIAIYYNVIIC 286

Qy     163 WSIFYFFKSFQYPLPWSEC-------------------------PVVR------------ 185
           :::|| | ||   |||  |                         | ::
Db     287 YTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVIGDHPKIQIKNSTFCMTAYP 346

Qy     186 NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAV 242
           | ::    ::  |   | :  || |   | ||  |   | : | :  || : | |   ::
Db     347 NLTMVNFTSQTNKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLAFCLFLAWVIVYASL 405

Qy     243 GKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVA 302
           |||:|||||:||:: ||||||   |:||: | ||  || :  |||  |: |  ||:: |
Db     406 AKGIKSSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKLTDATVWKDAA 465

Qy     303 TQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIM 362
           ||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:| |:|| ||
Db     466 TQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIFSVIGFMAN-- 523

Qy     363 NEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGL 422
             |  :||
Db     524 ERKVNIENV---------------------------------------------------- 532

Qy     423 DPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGIT 482
                 : || |:||: : ||:|  | ||||:::|||||: ||| :|  |:  |
Db     533 -----------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLDTMFATIETIV 581

Qy     483 TPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVIL 538
           | | |       :  | :||:|||:  |::|   : : | |   : | |:|:  | :| |
Db     582 TSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAASYALVIIAIF 641

Qy     539 ENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAY 598
           | : ::::|| ::| :::  |:||:|  |:   | ||:|  :  :   |  |    | |
Db     642 ELVGISYVYGLQRFCEDIEMMIGFKPNIFWKVCWAFVTPTILTFILCFSFYQW--EPMTY 699

Qy     599 SAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKA 658
           ::       ||   ||| | |   ::  : : ||::||:: :  |: |
Db     700 GSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIKMY--LAPG---------- 736

Qy     659 RMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETS 706
           | :         :  : : |  |   |  :  ||          || ||
Db     737 RFI---------ERLKLVCSPQPDWGPF-LAQHRGERYKNMIDPLGTS 774
```

```
RESULT 15
Q9BYZ7
ID   Q9BYZ7      PRELIMINARY;      PRT;   397 AA.
AC   Q9BYZ7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium channel-like protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Londono-Vallejo J.A.;
RT   "Sequence of a BAC carrying the entire hTERT gene.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY007685; AAG23290.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   397 AA;  43784 MW;  E70588E4D7FD9CF6 CRC64;

  Query Match             26.8%;  Score 1024;  DB 4;  Length 397;
  Best Local Similarity   38.1%;  Pred. No. 1.1e-74;
  Matches  189;  Conservative   87;  Mismatches  104;  Indels  116;  Gaps    5;

Qy      58 EDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
           ::|| |::| ||:|: |||:||||||||||||||  ||||:|:||:: |:  |||:| :|
Db      16 DERPKWDNKAQYLLSCIGFAVGLGNIWRFPYLCQTYGGGAFLIPYVIALVFEGIPIFHVE 75

Qy     118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
           ||:|||:|:||:|||  | | | |:|     :   : |||| |: | ::|   |||:|||
Db      76 LAIGQRLRKGSVGVWTAISPYLSGVGLGCVTLSFLISLYYNTIVAWVLWYLLNSFQHPLP 135

Qy     178 WSECP--VVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVW 235
           || ||  : | | |     ||: ||| :|||||: |:|:  |::|| : | : :||   |
Db     136 WSSCPPDLNRTGFV----EECQGSSAVSYFWYRQTLNITADINDSGSIQWWLLICLAASW 191

Qy     236 SIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDP 295
           ::  | | :||:::|||:||::|||||:||  ||:||| | ||  |::::|||     :|
Db     192 AVVYMCVIRGIETTGKVIYFTALFPYLVLTIFLIRGLTLPGATKGLIYLFTPN-----NP 246

Qy     296 QVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVL 355
           :|| : |||:|| | | ||| | |:|||   |:|  |  ::: :|  ||: |:: || ||
Db     247 RVWLDAATQIFFSLSLAFGGHIAFASYNSPRNDCQKDAVVIALVNRMTSLYASIAVFSVL 306

Qy     356 GFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKED 415
           |
Db     307 G------------------------------------------------------------ 307

Qy     416 QFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMI 475
                               | ||||: |||   | | :| |:::|| ||  ||| :|
Db     308 ------------------ASGPGLAFVVFTETDLHMPGAPVWAMLFFGMLFTLGLSTMF 348

Qy     476 GTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLI 535
           ||:  : ||:|                             ||||:: :||:::|:| | ::
Db     349 GTVEAVITPLLD-------------------------SGNYWLEIFDNFAASLNLLML 381

Qy     536 VILENIAVAWIYGPKK 551
             || : | ::|| |:
Db     382 AFLEVVGVVYVYGMKR 397


Search completed: April 21, 2003, 13:57:45
Job time : 101 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 13:46:57 ; Search time 27 Seconds
(without alignments)
1116.789 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 3817
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3610	94.6	727	1	NTT4_RAT	P31662 rattus norv
2	2522.5	66.1	730	1	NTT7_HUMAN	Q9h2j7 homo sapien
3	2494	65.3	729	1	NTT7_BOVIN	Q9xs59 bos taurus
4	2488	65.2	729	1	NTT7_RAT	Q08469 rattus norv
5	2184	57.2	439	1	NTT4_HUMAN	Q9h1v8 homo sapien
6	1147	30.0	225	1	NTT4_BOVIN	Q28001 bos taurus
7	1134	29.7	736	1	NTT5_HUMAN	Q9gzn6 homo sapien
8	1036.5	27.2	797	1	S6A5_HUMAN	Q9y345 homo sapien
9	1028	26.9	799	1	S6A5_RAT	P58295 rattus norv
10	1004	26.3	633	1	S6A9_RAT	P28572 rattus norv
11	995.5	26.1	692	1	S6A9_HUMAN	P48067 homo sapien
12	985.5	25.8	617	1	S6A2_HUMAN	P23975 homo sapien
13	985.5	25.8	638	1	S6A9_BOVIN	Q28039 bos taurus
14	981.5	25.7	619	1	S6A3_RAT	P23977 rattus norv
15	981.5	25.7	633	1	S6A9_MOUSE	P28571 mus musculu
16	976.5	25.6	619	1	S6A3_MOUSE	Q61327 mus musculu
17	975.5	25.6	617	1	S6A2_MOUSE	O55192 mus musculu
18	974.5	25.5	635	1	S6A8_RAT	P28570 rattus norv
19	969.5	25.4	614	1	S6AC_RAT	P48056 rattus norv
20	969.5	25.4	635	1	S6A8_RABIT	P31661 oryctolagus
21	968.5	25.4	614	1	S6AC_CANFA	P27799 canis famil
22	968.5	25.4	636	1	S6A7_HUMAN	Q99884 homo sapien
23	966.5	25.3	599	1	S6A1_RAT	P23978 rattus norv
24	965.5	25.3	620	1	S6A3_HUMAN	Q01959 homo sapien
25	964.5	25.3	620	1	S6A6_BOVIN	Q9mz34 bos taurus
26	963.5	25.2	635	1	S6A8_BOVIN	O18875 bos taurus
27	963.5	25.2	637	1	S6A7_RAT	P28573 rattus norv
28	963	25.2	635	1	S6A8_HUMAN	P48029 homo sapien
29	962.5	25.2	615	1	S6A2_BOVIN	P51143 bos taurus
30	958.5	25.1	614	1	S6AC_RABIT	P48055 oryctolagus
31	955	25.0	614	1	S6AC_HUMAN	P48065 homo sapien
32	951.5	24.9	599	1	S6A1_HUMAN	P30531 homo sapien
33	951.5	24.9	620	1	S6A6_HUMAN	P31641 homo sapien
34	951	24.9	602	1	S6AD_MOUSE	P31649 mus musculu
35	951	24.9	602	1	S6AD_RAT	P31646 rattus norv
36	948.5	24.8	614	1	S6AC_MOUSE	P31651 mus musculu
37	947.5	24.8	630	1	S6A4_CAVPO	O35899 cavia porce
38	946.5	24.8	630	1	S6A4_MOUSE	Q60857 mus musculu
39	946	24.8	598	1	S6A1_MUSCO	P48057 mus cookii
40	943.5	24.7	630	1	S6A4_RAT	P31652 rattus norv
41	941.5	24.7	620	1	S6A6_CANFA	Q00589 canis famil
42	941.5	24.7	630	1	S6A4_MACMU	Q9myx0 macaca mula
43	937.5	24.6	630	1	S6A4_HUMAN	P31645 homo sapien
44	937	24.5	693	1	S6A3_BOVIN	P27922 bos taurus
45	925.5	24.2	621	1	S6A6_RAT	P31643 rattus norv

ALIGNMENTS

RESULT 1
NTT4_RAT
ID NTT4_RAT STANDARD; PRT; 727 AA.
AC P31662;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Orphan sodium- and chloride-dependent neurotransmitter transporter
DE NTT4.
GN NTT4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93114444; PubMed=8093354;
RA Liu Q.-R., Mandiyan S., Lopez-Corcuera B., Nelson H., Nelson N.;
RT "A rat brain cDNA encoding the neurotransmitter transporter with an
RT unusual structure.";
RL FEBS Lett. 315:114-118(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94125086; PubMed=8294906;
RA el Mestikawy S., Giros B., Pohl M., Hamon M., Kingsmore S.F.,
RA Seldin M.F., Caron M.G.;
RT "Characterization of an atypical member of the Na+/Cl(-)-dependent
RT transporter family: chromosomal localization and distribution in
RT GABAergic and glutamatergic neurons in the rat brain.";
RL J. Neurochem. 62:445-455(1994).
CC -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CENTRAL NERVOUS
CC SYSTEM AND IS MORE ABUNDANT IN THE CEREBELLUM AND THE CEREBRAL
CC CORTEX.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNF).
CC ----------

CC ----------
DR EMBL; L06434; AAB24776.1; -.
DR EMBL; S68944; AAC60673.1; -.
DR PIR; S27043; S27043.
DR InterPro; IPR000175; Na/ntran_symport.
DR Pfam; PF00209; SNF; 1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/ntran_symport; 2.
DR PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.

```
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     68       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     69     89       1 (POTENTIAL).
FT   TRANSMEM     97    116       2 (POTENTIAL).
FT   TRANSMEM    141    161       3 (POTENTIAL).
FT   DOMAIN      162    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    243       4 (POTENTIAL).
FT   TRANSMEM    252    269       5 (POTENTIAL).
FT   TRANSMEM    305    322       6 (POTENTIAL).
FT   TRANSMEM    334    355       7 (POTENTIAL).
FT   DOMAIN      356    451       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    452    471       8 (POTENTIAL).
FT   TRANSMEM    495    513       9 (POTENTIAL).
FT   TRANSMEM    529    549       10 (POTENTIAL).
FT   TRANSMEM    570    591       11 (POTENTIAL).
FT   TRANSMEM    619    641       12 (POTENTIAL).
FT   DOMAIN      642    727       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    393    393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    261    261       Y -> C (IN REF. 2).
FT   CONFLICT    499    499       G -> S (IN REF. 2).
SQ   SEQUENCE   727 AA;  81055 MW;  C676048C0A6BDF7C CRC64;

  Query Match             94.6%;  Score 3610;  DB 1;  Length 727;
  Best Local Similarity   94.4%;  Pred. No. 5.5e-229;
  Matches  686;  Conservative   13;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
           ||||||||||||:|||||||||||||||||||||||||||||| |||||  |||||||||
Db       1 MPKNSKVTQREHSNEHVTESVADLLALEEPVDYKQSVLNVAGETGGKQKVAEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
           |||||||||||||:|||||||||||||||||||||||||||||:|||||||||||||||||
Db     121 GQRIRRGSIGVWHYVCPRLGGIGFSSCIVCLFVGLYYNVIIGWSVFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
           |||:|||:||||| |||||||||||||||||||:|||||||||||:|||| ||| ||
Db     181 CPVIRNGTVAVVEPECEKSSATTYFWYREALDISNSISESGGLNWKMTVCLLVAWSIVGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
           || ||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     241 AVVKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLDKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
           ||||| |||||||| |||||||||||||| ||||||||||||||||||| |||||||
Db     301 AATQVFFALGLGFGGVIAFSSYNKQDNNCHFDAALVSFINFFTSVLATLVVFAVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
           ||||||||||||||||:|||||||||||||||||||| || |||||||| |||||
Db     361 IMNEKCVVENAEKILGYLNSNVLSRDLIPPHVNFSHLTTKDYSEMYNVIMTVKEKQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
           ||||||| |||||||||||| | ||||||||||||||||||||||||:||||||
Db     481 ITTPIIDTFKVPKEMFTVGCCVFAFFVGLLFVQRSGNYFVTMFDDYSATLPLTVIVILEN 540

Qy     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
           |||||||| ||||||||||||||||||||||||||||||||||||||:|| |||
Db     541 IAVAWIYGTKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVSPPGYSA 600

Qy     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
           |||||||||||||| |||||||| ||||||||||:|||||||||||||||||| ||
Db     601 WIKEEAAERYLYFPNWAMALLITLIAVATLPIPVVFILRHFHLLSDGSNTLSVSYKKGRM 660

Qy     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
           |||||||||||||||||||||||||||||||||||||||||||||:| :||| || |||||
Db     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLESSSHPNGRYGSGYLLA 720

Qy     721 STPESEL 727
           |||||||
Db     721 STPESEL 727


RESULT 2
NTT7_HUMAN
ID   NTT7_HUMAN     STANDARD;      PRT;   730 AA.
AC   Q9H2J7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT73 (Orphan transporter v7-3).
GN   NTT73.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cerebellum;
RX   MEDLINE=20564208; PubMed=11112352;
RA   Farmer M.K., Robbins M.J., Medhurst A.D., Campbell D.A., Ellington K.,
RA   Duckworth M., Brown A.M., Middlemiss D.N., Price G.W., Pangalos M.N.;
RT   "Cloning and characterization of human NTT5 and v7-3: two orphan
RT   transporters of the Na(+)/Cl(-)-dependent neurotransmitter
RT   transporter gene family.";
RL   Genomics 70:241-252(2000).
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF265577; AAG41361.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     90       1 (POTENTIAL).
FT   TRANSMEM     98    117       2 (POTENTIAL).
FT   TRANSMEM    142    162       3 (POTENTIAL).
FT   DOMAIN      163    225       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    226    244       4 (POTENTIAL).
FT   TRANSMEM    253    270       5 (POTENTIAL).
FT   TRANSMEM    306    323       6 (POTENTIAL).
FT   TRANSMEM    335    356       7 (POTENTIAL).
FT   DOMAIN      357    452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    453    472       8 (POTENTIAL).
FT   TRANSMEM    496    514       9 (POTENTIAL).
FT   TRANSMEM    530    550       10 (POTENTIAL).
FT   TRANSMEM    571    592       11 (POTENTIAL).
FT   TRANSMEM    620    642       12 (POTENTIAL).
FT   DOMAIN      643    730       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    213    213       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    383    383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    394    394       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   730 AA;  81836 MW;  45963118E06CFFE6 CRC64;

  Query Match             66.1%;  Score 2522.5;  DB 1;  Length 730;
  Best Local Similarity   64.4%;  Pred. No. 8.3e-158;
  Matches  473;  Conservative  114;  Mismatches  135;  Indels   13;  Gaps    8;

Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
          ||||||| :||   : ||||| |||: |:  |  :| | | | |:   |   |||  : |
Db      1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy     59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
          | |||||||||||||:||||||||:|||||||||||||||:|||:||::|||||||||
Db     59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
          |:|||||||||||||:|| |:||||||:||:|| || ||||||||||:||| :||| |||
Db    119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy    178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237
          | :||:|:| |   || |||:|||||||:|||||||:|| ||||||||||||:|||  | :
Db    179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy    238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297
            :|: ||||||||::|||||||||||| |||:|  || |::||| |||||||  ||:|:|
Db    239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy    298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357
          ||| |||||| |||||||| |||||:|||||||  ||||||||||||||||| ||||
Db    299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy    358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417
          |||::|||| :|:| |: :|    :|:|:|| |:| | :| :||  : ::|  |||::|
Db    359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy    418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
           || |: | :|:||:|:|||||||||||||||||| ||||||||||:|||||| ||
Db    419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy    478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
          : || |||:|| || ||: || ||:  | :|||||||||||||||||||||| ::||
Db    479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy    538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
          |||||| ::||  |||::| :|||| | |:|:||||::||| :  |  ||:: :|:::||
Db    539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy    598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
          |:|||:::|:| :| :| | : : ::|:| | ||:||||::| |:|: | | |  ||:||
Db    599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy    657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
          : |::|:  || | |:|  |  |:||| |||      |     :  |:|:  ||| || |
Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727
          ||:|      |||:|
Db    716 YLMADIMPDMPESDL 730


RESULT 3
NTT7_BOVIN
ID   NTT7_BOVIN     STANDARD;      PRT;   729 AA.
AC   Q9XS59;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT73 (Orphan transporter v7-3).
GN   NTT73.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RA   Sakata K.;
RT   "Orphan transporter.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB020854; BAA77223.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     90       1 (POTENTIAL).
FT   TRANSMEM     98    117       2 (POTENTIAL).
FT   TRANSMEM    142    162       3 (POTENTIAL).
FT   DOMAIN      163    225       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    226    244       4 (POTENTIAL).
FT   TRANSMEM    253    270       5 (POTENTIAL).
FT   TRANSMEM    306    323       6 (POTENTIAL).
FT   TRANSMEM    335    356       7 (POTENTIAL).
FT   DOMAIN      357    452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    453    472       8 (POTENTIAL).
FT   TRANSMEM    496    514       9 (POTENTIAL).
FT   TRANSMEM    530    550       10 (POTENTIAL).
FT   TRANSMEM    571    592       11 (POTENTIAL).
FT   TRANSMEM    620    642       12 (POTENTIAL).
FT   DOMAIN      643    729       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    213    213       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    383    383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    394    394       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   729 AA;  81681 MW;  332FD83349C196A9 CRC64;

  Query Match             65.3%;  Score 2494;  DB 1;  Length 729;
  Best Local Similarity   64.0%;  Pred. No. 6e-156;
  Matches  471;  Conservative  115;  Mismatches  134;  Indels   16;  Gaps    9;

Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQS--VLNVAGEAGGKQKAVEEELD 56
          ||||||| :||   | | ||| |||: |:  |  :|:|  :::|  |   |  | | :
Db      1 MPKNSKVVKRELDDE-VIESVKDLLSNEDSADDAFKKSELIVDVPEE---KDTDVVERSE 56

Qy     57 AED-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFF 115
            :| |||||||||||||:||||||||:|||||||||||||||:|||:||::|||||||
Db     57 VKDARPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLLVIGIPLFF 116

Qy    116 LELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYP 175
          |||:|||||||||||||:|| |:||||||:||:|| || ||||||||||:||| :||| |
Db    117 LELSVGQRIRRGSIGVWNYISPQLGGIGFASCVVCFFVALYYNVIIGWSLFYFSQSFQQP 176

Qy    176 LPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVW 235
          ||| :||:|:| |   || |||||||||:|||||:|| ||||||||||||:|||  |
```

```
Db    177 LPWDQCPLVKNASHTFVEPECEKSSATTYYWYREALNISTSISESGGLNWKMTICLLAAW 236

Qy    236 SIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDP 295
          :  :|: ||||||||:||||||||||| |||:| ||| |:|||| |||||:|  ||:|
Db    237 VVVCLAMIKGIQSSGKIMYFSSLFPYVVLICFLIRALLLNGSVDGIRHMFTPELEIMLEP 296

Qy    296 QVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVL 355
          :|||| | |||| |||||||| ||||||:|||||||  ||||||||||:||||||| ||
Db    297 KVWREAAAQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSILATLVVFAVL 356

Qy    356 GFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKED 415
          |||||::|||||: ||:| |:  :    :|:|:|| |:||| :| :||  : ::|  |||:
Db    357 GFKANVINEKCIAENSEMIIKLVKMGNISQDIIPHHINFSAITAEDYDLIYDIIQKVKEE 416

Qy    416 QFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMI 475
          :| || |: | :||||:|:||||||||||||||||||| ||||||||||||:|||||||
Db    417 EFPALHLNACQIEDELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMF 476

Qy    476 GTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLI 535
          ||: || ||::|| || ||: || ||:  | :||:|||||||||||||||||||||| ::
Db    477 GTIEGIITPVVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIV 536

Qy    536 VILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTP 595
          ||||||||:::||  |||::| :|||| | |:|:||||::||| :  |  |||: :|::|
Db    537 VILENIAVSFVYGIDKFMEDLKDMLGFTPNRYYYYMWKYISPLMLLSLLIASIVNMGLSP 596

Qy    596 PAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVS 654
          | |:||::::|:|::| :| | | : |:|:|:| |||||||::|  :|: | | |  ||:
Db    597 PGYNAWMEDKASEKFLSYPTWGMVICISLMVLAILPIPVVFIIRRCNLIDDSSGNLASVT 656

Qy    655 YKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYG 714
          ||: |::|:  || | |:   |  |: ||  ||      |     :  |:|:  ||| ||
Db    657 YKRGRVLKEPVNL-EGDDASLIHGKISSEMSSPNFGKNIYRKQSGSPTLDTA--PNGRYG 713

Qy    715 RGYLLA---STPESEL 727
            |||:|     |||:|
Db    714 IGYLMADMPDMPESDL 729
```

```
RESULT 4
NTT7_RAT
ID   NTT7_RAT       STANDARD;      PRT;   729 AA.
AC   Q08469; Q63838;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT73 (Orphan transporter v7-3).
GN   NTT73.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93180651; PubMed=1363329;
RA   Uhl G.R., Kitayama S., Gregor P., Nanthakumar E., Persico A.M.,
RA   Shimada S.;
RT   "Neurotransmitter transporter family cDNAs in a rat midbrain library:
RT   'orphan transporters' suggest sizable structural variations.";
RL   Brain Res. Mol. Brain Res. 16:353-359(1992).
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L22022; AAA41729.1; -.
DR   EMBL; S56968; AAB25532.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     90       1 (POTENTIAL).
FT   TRANSMEM     98    117       2 (POTENTIAL).
FT   TRANSMEM    142    162       3 (POTENTIAL).
FT   DOMAIN      163    225       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    226    244       4 (POTENTIAL).
FT   TRANSMEM    253    270       5 (POTENTIAL).
FT   TRANSMEM    306    323       6 (POTENTIAL).
FT   TRANSMEM    335    356       7 (POTENTIAL).
FT   DOMAIN      357    452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    453    472       8 (POTENTIAL).
FT   TRANSMEM    496    514       9 (POTENTIAL).
FT   TRANSMEM    530    550       10 (POTENTIAL).
FT   TRANSMEM    571    592       11 (POTENTIAL).
FT   TRANSMEM    620    642       12 (POTENTIAL).
FT   DOMAIN      643    729       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    383    383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    394    394       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      1      3       MPK -> IP (IN REF. 1; AAB25532).
SQ   SEQUENCE   729 AA;  81596 MW;  ECDDC6403741B53E CRC64;
```

```
  Query Match             65.2%;  Score 2488;  DB 1;  Length 729;
  Best Local Similarity   63.9%;  Pred. No. 1.5e-155;
  Matches  473;  Conservative  109;  Mismatches  134;  Indels   24;  Gaps    8;

Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
          ||||||| :|:   : | ||| |||: |: |:   |:| | |  :        |:: |||
Db      1 MPKNSKVVKRD-LDDDVIESVKDLLSNEDSVEDVSKKSELIVDVQE-----TEKDTDAE 52

Qy     59 D-------RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111
          |       |||||||||||||:||||||||:|||||||||||||||:|||:||::|||
Db     53 DGSEVDDERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLLVIGI 112

Qy    112 PLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKS 171
          |||||||:|||||||||||||:|| |:||||||:||:|| || |||||||||::||| :|
Db    113 PLFFLELSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWTLFYFSQS 172

Qy    172 FQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCL 231
          || |||| :||:|:| |   :| ||||||||||:|||||| || ||||||||||||| ||
Db    173 FQQPLPWDQCPLVKNASHTYIEPECEKSSATTYYWYREALAISSSISESGGLNWKMTGCL 232

Qy    232 LVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVK 291
          |  | :  :|: ||||||||:|||||||||||| |||:| ||| |::||| |||||||
Db    233 LAAWVMVCLAMIKGIQSSGKIMYFSSLFPYVVLICFLIRSLLLNGSIDGIRHMFTPKLEM 292

Qy    292 MLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVV 351
          ||:|:|||| |||||| |||||||| |||||:|||||||  ||||||||||||||||||
Db    293 MLEPKVWREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVV 352

Qy    352 FVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMT 411
          | |||||||:|||| :|:| ||  | |  :| |:|| |:| | :| :||  : ::|
Db    353 FAVLGFKANIVNEKCISQNSEMILKLLKTGNVSWDVIPRHINLSAVTAEDYHVVYDIIQK 412

Qy    412 VKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGL 471
          |||::|: | |  | :||||:|:||||||||||||||||| ||||||||||||||||
Db    413 VKEEEFAVLHLKACQIEDELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGL 472

Qy    472 GSMIGTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLP 531
          ||| ||: || ||::|| || ||: || ||:  | :||:|||||||||||||||||||||
```

```
Db    473 GSMFGTIEGIITPVVDTFKVRKEILTVICCLLAFCIGLMFVQRSGNYFVTMFDDYSATLP 532

Qy    532 LTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQL 591
          | ::|||||||||:::||  ||:::||:|||| | ::|:||||::||| :  |  |||: :
Db    533 LLIVVILENIAVSFVYGIDKFLEDLTDMLGFAPSKYYYYMWKYISPLMLVTLLIASIVNM 592

Qy    592 GVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NT 650
          |::|| |:|||||:|:| :| :| | | :  :|:|:| ||:|||||:|  :|: | | |
Db    593 GLSPPGYNAWIKEKASEEFLSYPMWGMVVCFSLMVLAILPVPVVFVIRRCNLIDDSSGNL 652

Qy    651 LSVSYKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPN 710
           ||:||: |::|:  || : |:   |  |:||| ||      |     :  |:|:  ||
Db    653 ASVTYKRGRVLKEPVNL-DGDDASLIHGKIPSEMSSPNFGKNIYRKQSGSPTLDTA--PN 709

Qy    711 GPYGRGYLLA---STPESEL 727
          | || |||:|     |||:|
Db    710 GRYGIGYLMADMPDMPESDL 729


RESULT 5
NTT4_HUMAN
ID   NTT4_HUMAN     STANDARD;      PRT;   439 AA.
AC   Q9H1V8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT4 (Fragment).
GN   NTT4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Clark G.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL137790; CAC19682.1; ALT_INIT.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; PARTIAL.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; PARTIAL.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   NON_TER       1      1
FT   TRANSMEM     17     34       6 (POTENTIAL).
FT   TRANSMEM     46     67       7 (POTENTIAL).
FT   DOMAIN       68    163       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    164    183       8 (POTENTIAL).
FT   TRANSMEM    207    225       9 (POTENTIAL).
FT   TRANSMEM    241    261       10 (POTENTIAL).
FT   TRANSMEM    282    303       11 (POTENTIAL).
FT   TRANSMEM    331    353       12 (POTENTIAL).
FT   DOMAIN      354    439       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   439 AA;  49011 MW;  44634308C281C740 CRC64;
```

```
  Query Match             57.2%;  Score 2184;  DB 1;  Length 439;
  Best Local Similarity   95.9%;  Pred. No. 6.6e-136;
  Matches  421;  Conservative    1;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    289 LVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLAT 348
          | :|||||||| |||||| |||||||| |||||||||||||| ||||||||||||||
Db      1 LDQMLDPQVWREAATQVFFALGLGFGGVIAFSSYNKQDNNCHFDAALVSFINFFTSVLAT 60

Qy    349 LVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNV 408
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     61 LVVFAVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMYNV 120

Qy    409 IMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLIN 468
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    121 IMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLIN 180

Qy    469 LGLGSMIGTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
          |||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||||
Db    181 LGLGSMIGTMAGITTPIIDTFKVPKEMFTVGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA 240

Qy    529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    241 TLPLTLIVILENIAVAWIYGTKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 300

Qy    589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS 648
          |||||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    301 IQLGVTPPGYSAWIKEEAAERYLYFPNWAMALLITLIVVATLPIPVVFVLRHFHLLSDGS 360

Qy    649 NTLSVSYKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWN 708
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    361 NTLSVSYKKGRMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSGN 420

Qy    709 PNGPYGRGYLLASTPESEL 727
          ||| || ||||||||||||
Db    421 PNGRYGSGYLLASTPESEL 439


RESULT 6
NTT4_BOVIN
ID   NTT4_BOVIN     STANDARD;      PRT;   225 AA.
AC   Q28001;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT4 (Fragment).
GN   NTT4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RA   Jones E.M.C.;
RT   "Sodium-and chloride-dependent neurotransmitter transporters in bovine
RT   retina: identification and localization by in situ hybridization
RT   histochemistry.";
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; U19593; AAA61578.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   NON_TER       1      1
FT   TRANSMEM     <1      8       1 (POTENTIAL).
FT   TRANSMEM     16      35      2 (POTENTIAL).
FT   TRANSMEM     60      80      3 (POTENTIAL).
FT   DOMAIN       81      143     EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     144     162     4 (POTENTIAL).
FT   TRANSMEM     171     188     5 (POTENTIAL).
FT   TRANSMEM     224    >225     6 (POTENTIAL).
FT   CARBOHYD     105     105     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER      225     225
SQ   SEQUENCE   225 AA;  25374 MW;  3FA15D33F76D15C3 CRC64;

  Query Match             30.0%;  Score 1147;  DB 1;  Length 225;
  Best Local Similarity   93.8%;  Pred. No. 2.3e-68;
  Matches  211;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      82 NIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGG 141
           |:||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       1 NVWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYVCPRLGG 60

Qy     142 IGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSA 201
           |||||||||||||||||||||||||||||||||||||||||| |||:|||||||||||||
Db      61 IGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVSRNGTVAVVEAECEKSSA 120

Qy     202 TTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPY 261
           |||||||||||||:||||||||||||||||| | | |||| ||||||||||||||||||
Db     121 TTYFWYREALDISNSISESGGLNWKMTLCLLVAWRIVGMAVVKGIQSSGKVMYFSSLFPY 180

Qy     262 VVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVF 306
           |||||||||||||||:|||||||||||| |||||||||: |||:|
Db     181 VVLACFLVRGLLLRGAIDGILHMFTPKLDKMLDPQVWRDAATQIF 225


RESULT 7
NTT5_HUMAN
ID   NTT5_HUMAN     STANDARD;      PRT;   736 AA.
AC   Q9GZN6; Q9Y5I9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT5.
GN   NTT5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Sloan J.L., Mager S.;
RT   "Cloning of a new member of the neurotransmitter transporter family.";
RL   Abstr. - Soc. Neurosci. 24:1606-1606(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=20564208; PubMed=11112352;
RA   Farmer M.K., Robbins M.J., Medhurst A.D., Campbell D.A., Ellington K.,
RA   Duckworth M., Brown A.M., Middlemiss D.N., Price G.W., Pangalos M.N.;
RT   "Cloning and characterization of human NTT5 and v7-3: two orphan
RT   transporters of the Na(+)/Cl(-)-dependent neurotransmitter
RT   transporter gene family.";
RL   Genomics 70:241-252(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a catalog of human genes and proteins: sequencing and
RT   analysis of 500 novel complete protein coding human cDNAs.";
RL   Genome Res. 11:422-435(2001).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Highly expressed in peripheral tissues,
CC       particularly in testis, pancreas, and prostate.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF151977; AAD38044.1; -.
DR   EMBL; AF265578; AAG41362.1; -.
DR   EMBL; AL136856; CAB66790.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; FALSE_NEG.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     138       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    139     159       1 (POTENTIAL).
FT   TRANSMEM    177     197       2 (POTENTIAL).
FT   TRANSMEM    199     219       3 (POTENTIAL).
FT   DOMAIN      220     263       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    264     284       4 (POTENTIAL).
FT   TRANSMEM    290     310       5 (POTENTIAL).
FT   TRANSMEM    338     358       6 (POTENTIAL).
FT   TRANSMEM    383     403       7 (POTENTIAL).
FT   DOMAIN      404     495       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    496     516       8 (POTENTIAL).
FT   TRANSMEM    534     554       9 (POTENTIAL).
FT   TRANSMEM    568     588       10 (POTENTIAL).
FT   TRANSMEM    609     629       11 (POTENTIAL).
FT   TRANSMEM    659     679       12 (POTENTIAL).
FT   DOMAIN      680     736       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    229     229       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    181     181       W -> R (IN REF. 1).
FT   CONFLICT    307     307       G -> GG (IN REF. 1).
SQ   SEQUENCE   736 AA;  82199 MW;  E21EA16515446C35 CRC64;

  Query Match             29.7%;  Score 1134;  DB 1;  Length 736;
  Best Local Similarity   34.1%;  Pred. No. 6.3e-67;
  Matches  240;  Conservative  135;  Mismatches  287;  Indels   41;  Gaps   10;

Qy       5 SKVTQREHSSEHVTESVADLLALEEPVDYKQ-SVLNVAGEAGGKQKAVEEELDAED---- 59
           |  :: : |:  | |: |       :|   || |||      :   ||   |::   :
Db      44 SWTSEAQVSAARVAEAQA---RTSQP---KQISVLEALTASALNQKPTHEKVQMTEKKES 97

Qy      60 -----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLF 114
                || |:|| :|||||:|||:    :||| ||   :|| ::   |: :| ::|:||
Db      98 EVLLARPFWSSKTEYILAQVGFSMKPSCLWRFAYLWLNSGGCSFAAIYIFMLFLVGVPLL 157

Qy     115 FLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQY 174
           |||:| || :|:| :|||  | | :||:|:|| :||  :|||:||:  | |||  :|||:
Db     158 FLEMAAGQSMRQGGMGVWKIIAPWIGGVGYSSFMVCFILGLYFNVVNSWIIFYMSQSFQF 217
```

```
Qy     175 PLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVV 234
           |:|| :||:  | |    : |||::: : ||||::||  || | : |   : : |   :
Db     218 PVPWEKCPLTMNSS--GFDPECERTTPSIYFWYQQALKASDRIEDGGSPVYSLVLPFFLC 275

Qy     235 WSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLD 294
           | : |  :  |::|:|||:|   | |  ::  | :| ||| ||  |:  :   |:  : :
Db     276 WCLVGAFMINGLKSTGKVIYVLVLLPCFIIVGFFIRTLLLEGAKFGLQQLVVAKISDVYN 335

Qy     295 PQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVV 354
             ||     ||    |:| | |   :||  | |||  |  ||| ||  | :: |   | |
Db     336 MSVWSLAGGQVLSNTGIGLGSVASLASYMPQSNNCLSDAFLVSVINLLTLLVFTSFNFCV 395

Qy     355 LGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDY------MEMDNV 408
           ||| | ::  :|   ||| :|  :|   |  |  || ||  :  |  |       :
Db     396 LGFWATVITHRCCERNAEILLKLINLGKLPPDAKPP-VNLLYNPTSIYNAWLSGLPQHIK 454

Qy     409 IMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLIN 468
           | ::|       :  | :| :  |: :|   ||::| |||:  | | ||| :|||||:
Db     455 SMVLRE-------VTECNIETQFLKASEGPKFAFLSFVEAMSFLPPSVFWSFIFFLMLLA 507

Qy     469 LGLGSMIGTMAGITTPIIDT---SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDD 525
           :|| | || | || ||: ||    :  ::  ||  :  |: || | : ||:||: :  |
Db     508 MGLSSAIGIMQGIITPLQDTFSFFRKHTKLLIVGVFLLMFVCGLFFTRPSGSYFIRLLSD 567

Qy     526 YSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTT 585
           |    |: ::|: | :||:| || ::|: :|| :||      : ::|  : |: : ::
Db     568 YWIVFPIIVVVVFETMAVSWAYGARRFLADLTILLGHPISPIFGWLWPHLCPVVLLIIFV 627

Qy     586 ASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLL- 644
             :: | : |  | :|    : |    :| | : |:|||  :  ||||  ||    | :
Db     628 TMMVHLCMKPITYMSWDSSTSKEVLRPYPPWALLLMITLFAIVILPIPAYFVYCRIHRIP 687

Qy     645 -----SDGSNTLSVSYKKARMMKDISNLEENDETRFILSKVPS 682
                ||  | | |   :  :     ::: :  :   :| ||
Db     688 FRPKSGDGPMTASTSLPLSHQLTPSKEVQKEEILQVDETKYPS 730
```

RESULT 8
S6A5_HUMAN
ID S6A5_HUMAN STANDARD; PRT; 797 AA.
AC Q9Y345; O95288; Q9BX77;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sodium- and chloride-dependent glycine transporter 2 (GlyT2) (GlyT-2).
GN SLC6A5 OR GLYT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99059482; PubMed=9845349;
RA Morrow J.A., Collie I.T., Dunbar D.R., Walker G.B., Shahid M.,
RA Hill D.R.;
RT "Molecular cloning and functional expression of the human glycine
RT transporter GlyT2 and chromosomal localisation of the gene in the
RT human genome.";
RL FEBS Lett. 439:334-340(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spinal cord;
RX MEDLINE=99310916; PubMed=10381548;
RA Gallagher M.J., Burgess L.H., Brunden K.R.;
RT "Characterization of multiple forms of the human glycine transporter
RT type-2.";
RL Brain Res. Mol. Brain Res. 70:101-115(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Luyten W.;
RT "Cloning and expression of a human glycine transporter type II.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS ARG-184; ASN-463 AND ALA-751.
RC TISSUE=Spinal cord;
RX MEDLINE=20074920; PubMed=10606742;
RA Evans J., Herdon H., Cairns W., O'Brien E., Chapman C., Terrett J.,
RA Gloger I.;
RT "Cloning, functional characterisation and population analysis of a
RT variant form of the human glycine type 2 transporter.";
RL FEBS Lett. 463:301-306(1999).
CC -!- FUNCTION: Terminates the action of glycine by its high affinity
CC sodium-dependent reuptake into presynaptic terminals. May be
CC responsible for the termination of neurotransmission at
CC strychnine-sensitive glycinergic synapses.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in medulla, and to a lesser extent
CC in spinal cord and cerebellum.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNF).
CC --
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC --
DR EMBL; AF085412; AAC95145.1; -.
DR EMBL; AF142501; AAD27892.1; -.
DR EMBL; AF117999; AAK12641.1; -.
DR EMBL; AF352733; AAK29670.1; -.
DR Genew; HGNC:11051; SLC6A5.
DR MIM; 604159; -.
DR InterPro; IPR000175; Na/ntran_symport.
DR Pfam; PF00209; SNF; 1.
DR ProDom; PD000448; Na/ntran_symport; 2.
DR PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport; Polymorphism.

FT	DOMAIN	1	199	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	200	220	1 (POTENTIAL).
FT	TRANSMEM	228	247	2 (POTENTIAL).
FT	TRANSMEM	271	291	3 (POTENTIAL).
FT	DOMAIN	292	393	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	394	412	4 (POTENTIAL).
FT	TRANSMEM	421	438	5 (POTENTIAL).
FT	TRANSMEM	474	491	6 (POTENTIAL).
FT	TRANSMEM	503	524	7 (POTENTIAL).
FT	TRANSMEM	557	576	8 (POTENTIAL).
FT	TRANSMEM	604	622	9 (POTENTIAL).
FT	TRANSMEM	638	658	10 (POTENTIAL).
FT	TRANSMEM	679	698	11 (POTENTIAL).
FT	TRANSMEM	717	735	12 (POTENTIAL).
FT	DOMAIN	737	797	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	343	343	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	353	353	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	358	358	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	184	184	Q -> R.
FT				/FTId=VAR_011591.
FT	VARIANT	463	463	D -> N (IN 10% OF THE POPULATION).
FT				/FTId=VAR_011592.
FT	VARIANT	751	751	V -> A.
FT				/FTId=VAR_011593.
FT	CONFLICT	24	24	G -> S (IN REF. 2).
FT	CONFLICT	102	102	G -> S (IN REF. 2).
FT	CONFLICT	155	155	S -> G (IN REF. 2).
FT	CONFLICT	188	188	N -> D (IN REF. 2).
FT	CONFLICT	362	362	Q -> L (IN REF. 2).
FT	CONFLICT	582	582	T -> S (IN REF. 2).

SQ SEQUENCE 797 AA; 87359 MW; BBAA089AB7E8BEDA CRC64;

```
  Query Match             27.2%;  Score 1036.5;  DB 1;  Length 797;
  Best Local Similarity   34.0%;  Pred. No. 1.6e-60;
  Matches  223;  Conservative  111;  Mismatches  186;  Indels  135;  Gaps   14;

Qy      36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
           ||  ||    | | : ||          |  |:||| :||: :|::|||||:|||||  :
Db     173 SVATVATQEDEQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223

Qy      93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
           |||||:|:|||::| : |:|:|||||:::||   :| : ||  | | | | | :  |: :
Db     224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVL 282

Qy     153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC-------------------------PVVR-- 185
           : :|||||| :::|| | ||   |||  |                           | ::
Db     283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy     186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
                    | ::    ::  |   | :  || |   | ||  |   | : | : |||
Db     343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy     233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
           : | |   :: |||::||||:||:: ||||||   |:||: | ||  || :  |||  |:
Db     402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy     293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
           |  ||:: |||:|| |   :||:|  ||||| |||: |  :|:  |  ||: |  |:|
Db     462 TDATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy     353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
           |:|| ||    |  :||
Db     522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
                       : || |:||: : ||:|  | ||||:::|||||: |||
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy     473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
           :|  |:  | | | |      :  | :||:|||:  |::|   : : | |   : | |:|
Db     578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy     529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
           :  | :| | | : ::::|| ::| :::  |:||:|  |:  :| ||:|  :  :   |
Db     638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy     589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
           |   | | ::       ||  ||| | |   ::  : : ||::||::  ||
Db     698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 9
S6A5_RAT
ID   S6A5_RAT       STANDARD;      PRT;   799 AA.
AC   P58295;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Sodium-.and chloride-dependent glycine transporter 2 (GlyT2) (GlyT-2).
GN   SLC6A5 OR GLYT2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=94043045; PubMed=8226790;
RA   Liu Q.-R., Lopez-Corcuera B., Mandiyan S., Nelson H., Nelson N.;
RT   "Cloning and expression of a spinal cord- and brain-specific glycine
RT   transporter with novel structural features.";
RL   J. Biol. Chem. 268:22802-22808(1993).
CC   -!- FUNCTION: Terminates the action of glycine by its high affinity
CC       sodium-dependent reuptake into presynaptic terminals. May be
CC       responsible for the termination of neurotransmission at
CC       strychnine-sensitive glycinergic synapses.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Specifically expressed in spinal cord, brain
CC       stem, and to a lesser extent in the cerebellum.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L21672; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1    201       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    202    222       1 (POTENTIAL).
FT   TRANSMEM    230    249       2 (POTENTIAL).
FT   TRANSMEM    273    293       3 (POTENTIAL).
FT   DOMAIN      294    395       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    396    414       4 (POTENTIAL).
FT   TRANSMEM    423    440       5 (POTENTIAL).
FT   TRANSMEM    476    493       6 (POTENTIAL).
FT   TRANSMEM    505    526       7 (POTENTIAL).
FT   TRANSMEM    559    578       8 (POTENTIAL).
FT   TRANSMEM    606    624       9 (POTENTIAL).
FT   TRANSMEM    640    660       10 (POTENTIAL).
FT   TRANSMEM    681    700       11 (POTENTIAL).
FT   TRANSMEM    719    737       12 (POTENTIAL).
FT   DOMAIN      739    799       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    345    345       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    366    366       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   799 AA;  87908 MW;  9DBE239A7B071FCD CRC64;

  Query Match             26.9%;  Score 1028;  DB 1;  Length 799;
  Best Local Similarity   32.2%;  Pred. No. 6e-60;
  Matches  228;  Conservative  118;  Mismatches  208;  Indels  154;  Gaps   16;

Qy      43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
           | | : ||          |  |:||| :||: :|::|||||:|||||  :|||||:|:||
Db     185 EQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQNGGGAFLIPY 235

Qy     103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
           |::| : |:|:||||:::||   :| : ||  | | | | | :  |: : : :||||||
Db     236 LMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVLIAIYYNVIIC 294

Qy     163 WSIFYFFKSFQYPLPWSEC-------------------------PVVR------------ 185
           :::|| | ||   |||  |                         | ::
Db     295 YTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVIGDHPKIQIKNSTFCMTAYP 354

Qy     186 NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAV 242
           | ::    ::  |   | :  || |   | ||  |   | : | : || : | |   ::
Db     355 NLTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLPFCLFLAWVIVYASL 413

Qy     243 GKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVA 302
           |||::||||:||:: ||||||   |:||: | ||  || :  |||  |: |  ||:: |
Db     414 AKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKLTDATVWKDAA 473

Qy     303 TQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIM 362
           ||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:| |:|| ||
Db     474 TQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIFSVIGFMAN-- 531
```

```
Qy    363 NEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGL 422
Db    532 ERKVNIENV---------------------------------------------------- 540

Qy    423 DPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGIT 482
Db    541 -----------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLDTMFATIETIV 589

Qy    483 TPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVIL 538
Db    590 TSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAASYALVIIAIF 649

Qy    539 ENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAY 598
Db    650 ELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSFYQW--EPMTY 707

Qy    599 SAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKA 658
Db    708 GSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIKMY--LAPG----------- 744

Qy    659 RMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETS 706
Db    745 RFI---------ERLKLVCSPQPDWGPF-LAQHRGERYKNMIDPLGTS 782
```

RESULT 10
S6A9_RAT
ID S6A9_RAT STANDARD; PRT; 633 AA.
AC P28572;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sodium- and chloride-dependent glycine transporter 1 (GlyT1) (GlyT-1).
GN SLC6A9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92357797; PubMed=1353889;
RA Guastella J., Brecha N., Weigmann C., Lester H.A., Davidson N.;
RT "Cloning, expression, and localization of a rat brain high-affinity
RT glycine transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7189-7193(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92265306; PubMed=1534013;
RA Smith K.E., Borden L.A., Hartig P.R., Branchek T.A., Weinshank R.L.;
RT "Cloning and expression of a glycine transporter reveal
RT colocalization with NMDA receptors.";
RL Neuron 8:927-935(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93264081; PubMed=8494645;
RA Borowsky B., Mezey E., Hoffman B.J.;
RT "Two glycine transporter variants with distinct localization in the
RT CNS and peripheral tissues are encoded by a common gene.";
RL Neuron 10:851-863(1993).
CC -!- FUNCTION: Terminates the action of glycine by its high affinity
CC sodium-dependent reuptake into presynaptic terminals. May play a
CC role in regulation of glycine levels in NMDA receptor-mediated
CC neurotransmission.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/GLYT-1A/GLYT-2 (SHOWN HERE)
CC AND 2/GLYT-1B/GLYT-1; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: GLYT-1A IS FOUND IN THE GREY MATTER OF CNS AS
CC WELL AS IN MACROPHAGES AND MAST CELLS IN PERIPHERICAL TISSUES.
CC GLYT-1B IS ONLY FOUND IN THE WHITE MATTER OF CNS.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNF).
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; M95413; AAA41256.1; -.
DR EMBL; M88595; AAA41257.1; -.
DR EMBL; L13600; AAA73557.1; -.
DR PIR; JH0673; JH0673.
DR PIR; A46204; A46204.
DR InterPro; IPR000175; Na/ntran_symport.
DR Pfam; PF00209; SNF; 1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/ntran_symport; 2.
DR PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport; Amino-acid transport; Alternative splicing.

FT	DOMAIN	1	35 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	36	56 1 (POTENTIAL).
FT	TRANSMEM	64	83 2 (POTENTIAL).
FT	TRANSMEM	107	127 3 (POTENTIAL).
FT	DOMAIN	128	214 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	215	233 4 (POTENTIAL).
FT	TRANSMEM	242	259 5 (POTENTIAL).
FT	TRANSMEM	295	312 6 (POTENTIAL).
FT	TRANSMEM	324	345 7 (POTENTIAL).
FT	TRANSMEM	378	397 8 (POTENTIAL).
FT	TRANSMEM	426	444 9 (POTENTIAL).
FT	TRANSMEM	460	480 10 (POTENTIAL).
FT	TRANSMEM	501	520 11 (POTENTIAL).
FT	TRANSMEM	539	557 12 (POTENTIAL).
FT	DOMAIN	559	633 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	164	164 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	167	167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	177	177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	183	183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	1	10 MVGKGAKGML -> MAVAHGPVATSSPEQ (IN ISOFORM
FT			2).

SQ SEQUENCE 633 AA; 70570 MW; 7D25B555B97E9E17 CRC64;

Query Match 26.3%; Score 1004; DB 1; Length 633;
Best Local Similarity 31.4%; Pred. No. 1.7e-58;
Matches 224; Conservative 130; Mismatches 224; Indels 136; Gaps 16;

```
Qy     42 GEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGA 97
Db      5 GAKGMLNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGA 64

Qy     98 YLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYY 157
Db     65 FMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYY 123

Qy    158 NVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-VVE 193
Db    124 NVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHLFN 183

Qy    194 AECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVM 253
Db    184 YTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGKVV 243

Qy    254 YFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGF 313
Db    244 YFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGCAW 303

Qy    314 GGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEK 373
```

```
          ||:|  :|||| |||: |  ::|  |  ||| |  |:| :||| ||
Db    304 GGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN---------- 350

Qy    374 ILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDK 433
                                                      ||:|   : |
Db    351 ----------------------------------------------HLGVDVSRVAD---- 361

Qy    434 SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPK 493
            | ||||:|: ||:|  | || ||::|| ||| ||||:    :  : | |:|  :|
Db    362 --HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGN 417

Qy    494 EMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWI 546
          |         |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  ::: :|
Db    418 EWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIMYI 477

Qy    547 YGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEA 606
            || : : |::  |||| |  |:   |:|||  :  :   ::||    |  |:
Db    478 YGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN------- 528

Qy    607 AERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDI 664
            :  :| | :|:  |: | |  :|:  :| |      :||   |       :  :|
Db    529 ---HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSRDW 581

Qy    665 S-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
            | |:   |:     |:  |||       | |        :|   :| |: |
Db    582 GPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 631


RESULT 11
S6A9_HUMAN
ID   S6A9_HUMAN     STANDARD;      PRT;   692 AA.
AC   P48067;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium- and chloride-dependent glycine transporter 1 (GlyT1) (GlyT-1).
GN   SLC6A9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=94239375; PubMed=8183239;
RA   Kim K.-M., Kingsmore S.F., Han H., Yang-Feng T.L., Godinot N.,
RA   Seldin M.F., Caron M.G., Giros B.;
RT   "Cloning of the human glycine transporter type 1: molecular and
RT   pharmacological characterization of novel isoform variants and
RT   chromosomal localization of the gene in the human and mouse
RT   genomes.";
RL   Mol. Pharmacol. 45:608-617(1994).
CC   -!- FUNCTION: Terminates the action of glycine by its high affinity
CC       sodium-dependent reuptake into presynaptic terminals. May play a
CC       role in regulation of glycine levels in NMDA receptor-mediated
CC       neurotransmission.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; GLYT-1A, GLYT-1B AND GLYT-
CC       1C (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALL GLYT-1 SUBTYPES CAN BE FOUND IN BRAIN,
CC       KIDNEY, PANCREAS, LUNG, PLACENTA, LIVER. BUT GLYT-1C SUBTYPE IS
CC       ONLY FOUND IN THE BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S70609; AAB30784.1; -.
DR   EMBL; S70612; AAB30785.1; -.
DR   Genew; HGNC:11056; SLC6A9.
DR   MIM; 601019; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Amino-acid transport; Alternative splicing.
FT   DOMAIN        1     94       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     95    115       1 (POTENTIAL).
FT   TRANSMEM    123    142       2 (POTENTIAL).
FT   TRANSMEM    166    186       3 (POTENTIAL).
FT   DOMAIN      187    273       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    274    292       4 (POTENTIAL).
FT   TRANSMEM    301    318       5 (POTENTIAL).
FT   TRANSMEM    354    371       6 (POTENTIAL).
FT   TRANSMEM    383    404       7 (POTENTIAL).
FT   TRANSMEM    437    456       8 (POTENTIAL).
FT   TRANSMEM    485    503       9 (POTENTIAL).
FT   TRANSMEM    519    539       10 (POTENTIAL).
FT   TRANSMEM    560    579       11 (POTENTIAL).
FT   TRANSMEM    598    616       12 (POTENTIAL).
FT   DOMAIN      618    692       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    236    236       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     69       MAAAHGPVAPSSPEQVTLLPVQRSFFLPPFSGATPSTSLAE
FT                                SVLKVWHGAYNSGLLPQLMAQHSLAMAQ -> MVGKGAKGM
FT                                L (IN ISOFORM GLYT-1A).
FT   VARSPLIC     16     69       MISSING (IN ISOFORM GLYT-1B).
SQ   SEQUENCE   692 AA;  76823 MW;  FABA3243A0D98073 CRC64;

  Query Match             26.1%;  Score 995.5;  DB 1;  Length 692;
  Best Local Similarity   32.4%;  Pred. No. 6.8e-58;
  Matches  219;  Conservative  122;  Mismatches  210;  Indels  125;  Gaps   14;

Qy     50 AVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVL 105
          ||  |    |    |  | ::::::|  :|::|||||:||||||| :|||||:: || ::
Db     72 AVPSEATKRDQNLKRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGAFMFPYFIM 131

Qy    106 LIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSI 165
          ||  ||||||:||: ||   :| :|||  | |   |:|:   :|  ::|:||||:|  :
Db    132 LIFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYYNVVICIAF 190

Qy    166 FYFFKSFQYPLPWSEC--PVVRNGSVAVVEA-----------------------ECEKSSA 201
          :||| |  : |||: |  |    |::|                            :::|
Db    191 YYFFSSMTHVLPWAYCNNPWNTHDCAGVLDASNLTNGSRPAALPSNLSHLLNHSLQRTSP 250

Qy    202 TTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPY 261
          :  :|    | :|| |   | :   :  || | | :  : : :|::||||||:||:: |||
Db    251 SEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWLVVFLCLIRGVKSSGKVVYFTATFPY 310

Qy    262 VVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSS 321
          |||    |||: | || |||::  ||:  |:|: :|| : |:|:|: |   :||:|  :|
Db    311 VVLTILFVRGVTLEGAFDGIMYYLTPQWDKILEAKVWGDAASQIFYSLACAWGGLITMAS 370

Qy    322 YNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTN 381
          |||  |||: |  ::|  |  ||| |  |:| :||| ||              ¡
Db    371 YNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN---------------------- 409

Qy    382 VLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLA 441
                                                  ||:|   : |     | |||
Db    410 ----------------------------------HLGVDVSRVAD------HGPGLA 426

Qy    442 FIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPKEMF----- 496
          |:|: ||:|  | || ||::|| ||| ||||:    :  : | |:|  :|  |
Db    427 FVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGNEWILQKKT 484
```

```
Qy    497 --TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQ 554
            |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  :|: :||| : : |
Db    485 YVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVAIMYIYGHRNYFQ 544

Qy    555 ELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFP 614
          ::  |||| |  |:   |:|||| :  :   ::||    |  |:          :  :|
Db    545 DIQMMLGFPPPLFFQICWRFVSPAIIFFILVFTVIQ--YQPITYN---------HYQYP 592

Qy    615 NWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDIS-NLEEND 671
          | :|:  |: |  |  :|:  :| |      :||   |       :  :|    | |:
Db    593 GWAVAIGFLMALSSVLCIPLYAMFRL----CRTDGDTLLQRLKNATKPSRDWGPALLEHR 648

Qy    672 ETRFILSKVPSEAPSP 687
            |:     |: ||||
Db    649 TGRY----APTIAPSP 660
```

RESULT 12
S6A2_HUMAN
ID S6A2_HUMAN STANDARD; PRT; 617 AA.
AC P23975;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sodium-dependent noradrenaline transporter (Norepinephrine
DE transporter) (NET).
GN SLC6A2 OR NAT1 OR NET1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91179515; PubMed=2008212;
RA Pacholczyk T., Blakely R.D., Amara S.G.;
RT "Expression cloning of a cocaine- and antidepressant-sensitive human
RT noradrenaline transporter.";
RL Nature 350:350-354(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=96028030; PubMed=7488042;
RA Poerzgen P., Boenisch H., Bruss M.;
RT "Molecular cloning and organization of the coding region of the human
RT norepinephrine transporter gene.";
RL Biochem. Biophys. Res. Commun. 215:1145-1150(1995).
RN [3]
RP VARIANT OI PRO-457.
RX MEDLINE=20132303; PubMed=10684912;
RA Shannon J.R., Flattem N.L., Jordan J., Jacob G., Black B.K.,
RA Biaggioni I., Blakely R.D., Robertson D.;
RT "Orthostatic intolerance and tachycardia associated with
RT norepinephrine-transporter deficiency.";
RL New Engl. J. Med. 342:541-549(2000).
CC -!- FUNCTION: TERMINATES THE ACTION OF NORADRENALINE BY ITS HIGH
CC   AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DISEASE: DEFECTS IN SLC6A2 ARE A CAUSE OF ORTHOSTATIC INTOLERANCE
CC   (OI), A SYNDROME CHARACTERIZED BY LIGHTHEADEDNESS, FATIGUE,
CC   ALTERED MENTATION AND SYNCOPE. IT IS ASSOCIATED WITH POSTURAL
CC   TACHYCARDIA. PLASMA NOREPINEPHRINE CONCENTRATION IS ABNORMALLY
CC   HIGH.
CC -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
CC   STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC   FAMILY (SNF).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M65105; AAA59943.1; -.
DR EMBL; X91117; CAA62566.1; -.
DR EMBL; X91120; CAA62566.1; JOINED.
DR EMBL; X91121; CAA62566.1; JOINED.
DR EMBL; X91122; CAA62566.1; JOINED.
DR EMBL; X91123; CAA62566.1; JOINED.
DR EMBL; X91124; CAA62566.1; JOINED.
DR EMBL; X91125; CAA62566.1; JOINED.
DR EMBL; X91126; CAA62566.1; JOINED.
DR EMBL; X91127; CAA62566.1; JOINED.
DR EMBL; X91118; CAA62566.1; JOINED.
DR EMBL; X91119; CAA62566.1; JOINED.
DR PIR; S14278; S14278.
DR Genew; HGNC:11048; SLC6A2.
DR MIM; 163970; -.
DR MIM; 604715; -.
DR InterPro; IPR000175; Na/ntran_symport.
DR Pfam; PF00209; SNF; 1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/ntran_symport; 1.
DR PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport; Disease mutation; Polymorphism.

FT	DOMAIN	1	64	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	65	85	1 (POTENTIAL).
FT	TRANSMEM	93	112	2 (POTENTIAL).
FT	TRANSMEM	136	156	3 (POTENTIAL).
FT	DOMAIN	157	234	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	235	253	4 (POTENTIAL).
FT	TRANSMEM	262	279	5 (POTENTIAL).
FT	TRANSMEM	315	332	6 (POTENTIAL).
FT	TRANSMEM	344	365	7 (POTENTIAL).
FT	TRANSMEM	398	417	8 (POTENTIAL).
FT	TRANSMEM	444	462	9 (POTENTIAL).
FT	TRANSMEM	478	498	10 (POTENTIAL).
FT	TRANSMEM	519	538	11 (POTENTIAL).
FT	TRANSMEM	557	575	12 (POTENTIAL).
FT	DOMAIN	576	617	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	184	184	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	192	192	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	198	198	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	69	69	V -> I (IN DBSNP:1805064).
FT				/FTId=VAR_011756.
FT	VARIANT	99	99	T -> I (IN DBSNP:1805065).
FT				/FTId=VAR_011757.
FT	VARIANT	245	245	V -> I (IN DBSNP:1805066).
FT				/FTId=VAR_011758.
FT	VARIANT	292	292	N -> T (IN DBSNP:5563).
FT				/FTId=VAR_011759.
FT	VARIANT	356	356	V -> L (IN DBSNP:5565).
FT				/FTId=VAR_011760.
FT	VARIANT	369	369	A -> P (IN DBSNP:5566).
FT				/FTId=VAR_011761.
FT	VARIANT	375	375	N -> S (IN DBSNP:5567).
FT				/FTId=VAR_011762.
FT	VARIANT	457	457	A -> P (IN OI; LOSS OF FUNCTION).
FT				/FTId=VAR_010022.
FT	VARIANT	463	463	K -> R (IN DBSNP:5570).
FT				/FTId=VAR_011763.
FT	VARIANT	478	478	G -> S (IN DBSNP:1805067).
FT				/FTId=VAR_011764.
FT	VARIANT	528	528	F -> C (IN DBSNP:5558).
FT				/FTId=VAR_011765.
FT	VARIANT	548	548	Y -> H (IN DBSNP:5559).
FT				/FTId=VAR_011766.

SQ SEQUENCE 617 AA; 69332 MW; BDC6DF31316907BB CRC64;

Query Match 25.8%; Score 985.5; DB 1; Length 617;

```
  Best Local Similarity   34.5%;  Pred. No. 2.7e-57;
  Matches  219;  Conservative  110;  Mismatches  200;  Indels  105;  Gaps   14;

Qy      28 EEPVDYKQSV-LNVAGEAGGKQ-KAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWR 85
           |:|:  :::  | |  |  | |      : ||: |  |  |: ::|: :|||:| | |:||
Db      22 EQPLRARKTAELLVVKERNGVQCLLAPRDGDAQPRETWGKKIDFLLSVVGFAVDLANVWR 81

Qy      86 FPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFS 145
           ||||| ||||||:|:|| : ||| |:|||::|||:||  | |:  ||  |||   |:|::
Db      82 FPYLCYKNGGGAFLIPYTLFLIIAGMPLFYMELALGQYNREGAATVWK-ICPFFKGVGYA 140

Qy     146 SCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSEC----------PVVRNGSVAVVEA 194
             :: |:|| |||||| ||::| | ||   |||::|           | : ||||
Db     141 VILIALYVGFYYNVIIAWSLYYLFSSFTLNLPWTDCGHTWNSPNCTDPKLLNGSVLGNHT 200

Qy     195 ECEKSSAT--TYFWYREALDISDS--ISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSG 250
           : |  |    |: |  | : :|  | : |   |:: |||:||  :   :: ||:::||
Db     201 KYSKYKFTPAAEFYERGVLHLHESSGIHDIGLPQWQLLLCLMVVVIVLYFSLWKGVKTSG 260

Qy     251 KVMYFSSLFPYVVLACFLVRGLLLRGAVDGI---LHMFTPKLVKMLDPQVWREVATQVFF 307
           ||:: ::  || ||   || |: | || :||   ||:      :: :  || : |||:||
Db     261 KVVWITATLPYFVLFVLLVHGVTLPGASNGINAYLHI---DFYRLKEATVWIDAATQIFF 317

Qy     308 GLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCV 367
            || ||| :| |:|||| ||||: |  | | || ||  :| :||: |:    |
Db     318 SLGAGFGVLIAFASYNKFDNNCYRDALLTSSINCITSFVSGFAIFSILGYMAH--EHKVN 375

Qy     368 VENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLL 427
           :|
Db     376 IE---------------------------------------------------------- 377

Qy     428 EDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID 487
             | : :| || || : ||::    | ||:|:||:||: ||| | :| |  : | : |
Db     378 ----DVATEGAGLVFILYPEAISTLSGSTFWAVVFFVMLLALGLDSSMGGMEAVITGLAD 433

Qy     488 TSKVPK---EMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
            :| |   ::|| |    |||: |  : : | | :|: | ::|   :   |::| | |:
Db     434 DFQVLKRHRKLFTFGVTFSTFLLALFCITKGGIYVLTLLDTFAAGTSILFAVLMEAIGVS 493

Qy     545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
           | ||  :|  :: :|:|||  ::   ||||||  :  :   |||     |  |  :|
Db     494 WFYGVDRFSNDIQQMMGFRPGLYWRLCWKFVSPAFLLFVVVVSIINF--KPLTYDDYI-- 549

Qy     605 EAAERYLYFPNWP--MALLITLIVVATLPIPVVF 636
                 || |   :   | |  :  :|| |::
Db     550 --------FPPWANWVGWGIALSSMVLVPIYVIY 575


RESULT 13
S6A9_BOVIN
ID   S6A9_BOVIN     STANDARD;      PRT;   638 AA.
AC   Q28039; Q28040; Q29420;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Sodium- and chloride-dependent glycine transporter 1 (GlyT1) (GlyT-1).
GN   SLC6A9.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RA   Jones E.M.C.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Terminates the action of glycine by its high affinity
CC       sodium-dependent reuptake into presynaptic terminals. May play a
CC       role in regulation of glycine levels in NMDA receptor-mediated
CC       neurotransmission (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 3 isoforms; 1/GLYT-1B (shown here),
CC       2/GLYT-1A and 3/GLYT-1F; are produced by alternative splicing.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U52687; AAB01158.1; -.
DR   EMBL; U52688; AAB01159.1; -.
DR   EMBL; U52689; AAB01160.1; -.
DR   EMBL; U52690; AAB01161.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Amino-acid transport; Alternative splicing.
FT   DOMAIN        1     40       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     41     61       1 (POTENTIAL).
FT   TRANSMEM     69     88       2 (POTENTIAL).
FT   TRANSMEM    112    132       3 (POTENTIAL).
FT   DOMAIN      133    219       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    220    238       4 (POTENTIAL).
FT   TRANSMEM    247    264       5 (POTENTIAL).
FT   TRANSMEM    300    317       6 (POTENTIAL).
FT   TRANSMEM    329    350       7 (POTENTIAL).
FT   TRANSMEM    383    402       8 (POTENTIAL).
FT   TRANSMEM    431    449       9 (POTENTIAL).
FT   TRANSMEM    465    485       10 (POTENTIAL).
FT   TRANSMEM    506    525       11 (POTENTIAL).
FT   TRANSMEM    544    562       12 (POTENTIAL).
FT   DOMAIN      564    638       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    182    182       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     14       MAAAQGPVAPSSLE -> MVGKGAKGM (IN ISOFORM
FT                                2).
FT   VARSPLIC    537    638       YPSWRVRIGFLMALSSVICIPLYALFQFCRTDGDTLLHRLK
FT                                NATKPSRDWGPALLEHRTRRYAPTTTPSPEDGLEVQPLHPD
FT                                KAQIPMVGSNGSSRFQDSRI -> SSQTGLPLFTCQIAPAH
FT                                VPQPLSGARTPSPKPWSVRVSVLRAPLCSDSPGRAASNPL
FT                                (IN ISOFORM 3).
SQ   SEQUENCE   638 AA;  71318 MW;  97ADDBF28EC2A443 CRC64;

  Query Match             25.8%;  Score 985.5;  DB 1;  Length 638;
  Best Local Similarity   31.6%;  Pred. No. 2.8e-57;
  Matches  217;  Conservative  124;  Mismatches  214;  Indels  131;  Gaps   15;

Qy      50 AVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVL 105
           ||  |   :|    |  | ::::::|  :|::|||||:|||||||  :|||||:: || ::
Db      18 AVPSEATKKDQNLKRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGAFMFPYFIM 77

Qy     106 LIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSI 165
           ||  ||||||:||: ||   :| :|||  | |   |:|:   :|  ::|:||||:|  :
Db      78 LIFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYYNVVICIAF 136

Qy     166 FYFFKSFQYPLPWSEC--------------PVVRNGS---------VAVVEAECEKSSA 201
           :||| |    |||: |               | : |||           :    :::|
Db     137 YYFFSSMTPVLPWTYCNNPWNTPDCMSVLDNPNITNGSQPPALPGNVSQALNQTLKRTSP 196

Qy     202 TTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPY 261
           :  :|    | :|| |   | :   :  || | | :  : : :|::|||||:||:: |||
Db     197 SEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGKVVYFTATFPY 256

Qy     262 VVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSS 321
```

```
Db    257 VVLTILFIRGVTLEGAFTGIMYYLTPQWDKILEAKVWGDAASQIFYSLGCAWGGLVTMAS 316

Qy    322 YNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTN 381

Db    317 YNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN---------------------- 355

Qy    382 VLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLA 441

Db    356 ------------------------------------HLGVDVSRVAD------HGPGLA 372

Qy    442 FIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPKEMF----- 496

Db    373 FVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGNEWILQKKT 430

Qy    497 --TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQ 554

Db    431 YVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVIISCIMCVSIMYIYGHQNYFQ 490

Qy    555 ELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFP 614

Db    491 DIQMMLGFPPPLFFQICWRFVSPAIIFFILIFSVIQ--YQPITYNQY----------QYP 538

Qy    615 NWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDIS-NLEEND 671

Db    539 SWRVRIGFLMALSSVICIPLYALFQF----CRTDGDTLLHRLKNATKPSRDWGPALLEHR 594

Qy    672 ETRFILSKVPSEAPSP------MPTH 691

Db    595 TRRY----APTTTPSPEDGLEVQPLH 616


RESULT 14
S6A3_RAT
ID   S6A3_RAT       STANDARD;      PRT;   619 AA.
AC   P23977;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium-dependent dopamine transporter (DA transporter) (DAT).
GN   SLC6A3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92054540; PubMed=1948035;
RA   Kilty J.E., Lorang D., Amara S.G.;
RT   "Cloning and expression of a cocaine-sensitive rat dopamine
RT   transporter.";
RL   Science 254:578-579(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92054539; PubMed=1948034;
RA   Shimada S., Kitayama S., Lin C.-L., Patel A., Nanthakumar E.,
RA   Gregor P., Kuhar M., Uhl G.;
RT   "Cloning and expression of a cocaine-sensitive dopamine transporter
RT   complementary DNA.";
RL   Science 254:576-578(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92111727; PubMed=1765147;
RA   Giros B., el Mestikawy S., Bertrand L., Caron M.G.;
RT   "Cloning and functional characterization of a cocaine-sensitive
RT   dopamine transporter.";
RL   FEBS Lett. 295:149-154(1991).
RN   [4]
RP   MUTAGENESIS.
RX   MEDLINE=92366546; PubMed=1502198;
RA   Kitayama S., Shimada S., Xu H., Markham L., Donovan D.M., Uhl G.;
RT   "Dopamine transporter site-directed mutations differentially alter
RT   substrate transport and cocaine binding.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:7782-7785(1992).
CC   -!- FUNCTION: TERMINATES THE ACTION OF DOPAMINE BY ITS HIGH
CC       AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
CC       STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M80233; AAA41100.1; -.
DR   EMBL; M80570; AAA73143.1; -.
DR   EMBL; S76145; AAB21099.1; -.
DR   PIR; S20346; S20346.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     68       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     69     89       1 (POTENTIAL).
FT   TRANSMEM     96    116       2 (POTENTIAL).
FT   TRANSMEM    140    160       3 (POTENTIAL).
FT   DOMAIN      161    236       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    237    255       4 (POTENTIAL).
FT   TRANSMEM    264    281       5 (POTENTIAL).
FT   TRANSMEM    317    334       6 (POTENTIAL).
FT   TRANSMEM    346    367       7 (POTENTIAL).
FT   TRANSMEM    400    419       8 (POTENTIAL).
FT   TRANSMEM    446    464       9 (POTENTIAL).
FT   TRANSMEM    480    500       10 (POTENTIAL).
FT   TRANSMEM    521    540       11 (POTENTIAL).
FT   TRANSMEM    559    577       12 (POTENTIAL).
FT   DOMAIN      578    619       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    181    181       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    196    196       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    597    597       E -> K (IN REF. 3).
SQ   SEQUENCE   619 AA;  68746 MW;  DF3A30C981095D24 CRC64;

  Query Match             25.7%;  Score 981.5;  DB 1;  Length 619;
  Best Local Similarity   35.6%;  Pred. No. 4.9e-57;
  Matches  213;  Conservative  100;  Mismatches  186;  Indels   99;  Gaps   11;

Qy     55 LDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLF 114

Db     55 VEAQERETWSKKIDFLLSVIGFAVDLANVWRFPYLCYKNGGGAFLVPYLLFMVIAGMPLF 114

Qy    115 FLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQY 174

Db    115 YMELALGQFNREGAAGVWK-ICPVLKGVGFTVILISFYVGFFYNVIIAWALHYFFSSFTM 173

Qy    175 PLPWSECPVVRNGSVAVVEAECEKSS----------ATTYFWYREALDISDS--ISE 219

Db    174 DLPWIHCNNTWN-SPNCSDAHASNSSDGLGLNDTFGTTPAAEYF-ERGVLHLHQSRGIDD 231

Qy    220 SGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVD 279

Db    232 LGPPRWQLTACLVLVIVLLYFSLWKGVKTSGKVVWITATMPYVVLTALLLRGVTLPGAMD 291

Qy    280 GILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFI 339
```

```
Db    292 GIRAYLSVDFYRLCEASVWIDAATQVCFSLGVGFGVLIAFSSYNKFTNNCYRDAIITTSI 351

Qy    340 NFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTT 399

Db    352 NSLTSFSSGFVVFSFLGYMAQ-----------------KHNVPIRDV------------ 381

Qy    400 KDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWS 459

Db    382 ---------------------------------ATDGPGLIFIIYPEAIATLPLSSAWA 407

Qy    460 VMFFLMLINLGLGSMIGTMAGITTPIIDTSKV---PKEMFTVGCCVFTFLVGLLFVQRSG 516

Db    408 AVFFLMLLTLGIDSAMGGMESVITGLVDEFQLLHRHRELFTLGIVLATFLLSLFCVTNGG 467

Qy    517 NYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVS 576

Db    468 IYVFTLLDHFAAGTSILFGVLIEAIGVAWFYGVQQFSDDIKQMTGQRPNLYWRLCWKLVS 527

Qy    577 PLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMAL--LITLIVVATLPI 632

Db    528 PCFLLYVVVVSIVTF--RPPHYGAYI----------FPDWANALGWIIATSSMAMVPI 573


RESULT 15
S6A9_MOUSE
ID   S6A9_MOUSE     STANDARD;      PRT;   633 AA.
AC   P28571;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium- and chloride-dependent glycine transporter 1 (GlyT1) (GlyT-1).
GN   SLC6A9 OR GLYT1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92316254; PubMed=1618338;
RA   Liu Q.-R., Nelson H., Mandiyan S., Lopez-Corcuera B., Nelson N.;
RT   "Cloning and expression of a glycine transporter from mouse brain.";
RL   FEBS Lett. 305:110-114(1992).
CC   -!- FUNCTION: Terminates the action of glycine by its high affinity
CC       sodium-dependent reuptake into presynaptic terminals. May play a
CC       role in regulation of glycine levels in NMDA receptor-mediated
CC       neurotransmission.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; GLYT-1A (SHOWN HERE) AND GLYT-
CC       1B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X67056; CAA47440.1; -.
DR   PIR; S23151; S23151.
DR   MGD; MGI:95760; Glyt1.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Amino-acid transport; Alternative splicing.
FT   DOMAIN        1     35       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     36     56       1 (POTENTIAL).
FT   TRANSMEM     64     83       2 (POTENTIAL).
FT   TRANSMEM    107    127       3 (POTENTIAL).
FT   DOMAIN      128    214       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    215    233       4 (POTENTIAL).
FT   TRANSMEM    242    259       5 (POTENTIAL).
FT   TRANSMEM    295    312       6 (POTENTIAL).
FT   TRANSMEM    324    345       7 (POTENTIAL).
FT   TRANSMEM    378    397       8 (POTENTIAL).
FT   TRANSMEM    426    444       9 (POTENTIAL).
FT   TRANSMEM    460    480       10 (POTENTIAL).
FT   TRANSMEM    501    520       11 (POTENTIAL).
FT   TRANSMEM    539    557       12 (POTENTIAL).
FT   DOMAIN      559    633       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    164    164       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    183    183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     10       MVGKGAKGML -> MASAQGPVATPSPEQ (IN ISOFORM
FT                                GLYT-1B).
SQ   SEQUENCE   633 AA;  70775 MW;  FD64AA561F470793 CRC64;

  Query Match             25.7%;  Score 981.5;  DB 1;  Length 633;
  Best Local Similarity   31.3%;  Pred. No. 5e-57;
  Matches  217;  Conservative  126;  Mismatches  220;  Indels  131;  Gaps   15;

Qy     42 GEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGA 97

Db      5 GAKGMLNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGA 64

Qy     98 YLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYY 157

Db     65 FMFPYFIMLIFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYY 123

Qy    158 NVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-VVE 193

Db    124 NVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPAALSGNLSHLFN 183

Qy    194 AECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVM 253

Db    184 YTLQRTSPSEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGKVV 243

Qy    254 YFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGF 313

Db    244 YFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPQWDKYLEAKVWGDAASQIFYSLGCAW 303

Qy    314 GGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEK 373

Db    304 GGLITMASYNKFHNNCYRDSVIISITNCATRLYAGFVIFSILGFMAN------------ 350

Qy    374 ILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDK 433

Db    351 ---------------------------------------------HLGVDVSRVAD---- 361

Qy    434 SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPK 493

Db    362 --HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGN 417

Qy    494 EMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWI 546

Db    418 EWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIMYI 477

Qy    547 YGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEA 606

Db    478 YGHRNYFQDIQMMLGFPPPLFFQICWRFVSPAIIFFILIFTVIQ--YRPITYN------- 528

Qy    607 AERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDI 664

Db    529 ---HYQYPGWAVRIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSRDW 581

Qy    665 S-NLEENDETRFILSKVPSEAPSP------MPTH 691
```

```
                 | |:   |:     |:  |||       | |
Db     582 GPALLEHRTGRY----APTTTPSPEDGFEVQPLH 611
```

Search completed: April 21, 2003, 13:56:00
Job time : 30 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 13:53:57 ; Search time 48 Seconds
(without alignments)
1456.036 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 3817
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3610	94.6	727	2	S27043	neurotransmitter t
2	3595	94.2	727	2	I56506	Na+/Cl(-)-dependen
3	2483	65.1	730	2	I52632	sodium-dependent n
4	2463	64.5	728	2	I65413	sodium-dependent n
5	1302.5	34.1	616	2	S50998	transporter protei
6	1028	26.9	799	2	A48716	glycine transporte
7	1004	26.3	633	2	I58140	glycine transporte
8	1003	26.3	638	2	JH0673	glycine transport
9	997.5	26.1	638	2	I77912	glycine transporte
10	995.5	26.1	692	2	I57956	glycine transporte
11	985.5	25.8	617	2	S14278	noradrenaline tran
12	981.5	25.7	619	2	S20346	dopamine transport
13	981.5	25.7	619	2	I59558	dopamine transport
14	981.5	25.7	633	2	S23151	glycine transporte
15	975.5	25.6	635	2	A46061	Na(+)-dependent cr
16	974.5	25.5	635	2	S23431	choline transport
17	973.5	25.5	635	2	JC2386	creatine transport
18	968.5	25.4	614	2	A41757	betaine transport
19	966.5	25.3	599	1	ACRTGT	gamma-aminobutyric
20	965.5	25.3	620	2	A48980	dopamine transport
21	965	25.3	620	2	I57937	dopamine transport
22	963.5	25.2	637	2	JH0674	L-proline transpor
23	963	25.2	635	2	G02095	creatine transport
24	962.5	25.2	615	2	S43285	noradrenaline tran
25	957.5	25.1	602	2	I55651	noradrenaline tran
26	955.5	25.0	620	2	S46487	taurine transporte
27	955	25.0	614	2	S68236	betaine/GABA trans
28	953	25.0	611	2	S46260	creatine transport
29	951.5	24.9	599	2	S11073	gamma-aminobutyric
30	951.5	24.9	620	2	G01426	taurine transporte
31	951	24.9	602	2	A44409	gamma-aminobutyric
32	951	24.9	602	2	A45078	gamma-aminobutyric
33	948.5	24.8	614	2	A43390	gamma-aminobutyric
34	947	24.8	607	2	S19585	serotonin transpor
35	943.5	24.7	630	2	S30604	neurotransmitter t
36	941.5	24.7	655	2	A46270	Na and Cl dependen
37	937.5	24.6	630	2	A47398	serotonin transpor
38	937	24.5	693	2	A41617	dopamine transport
39	928.5	24.3	619	2	S29839	taurine transport
40	925.5	24.2	621	2	I57939	taurine transporte
41	924	24.2	598	2	F46027	gamma-aminobutyric
42	923.5	24.2	627	2	JH0695	gamma-aminobutyric
43	917	24.0	598	2	I51368	gamma-aminobutyric
44	915.5	24.0	627	2	B44409	gamma-aminobutyric
45	914	23.9	597	2	S65673	GABA transport pro

ALIGNMENTS

RESULT 1
S27043
neurotransmitter transport protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S27043
R;Liu, Q.R.; Mandiyan, S.; Lopez-Corcuera, B.; Nelson, H.; Nelson, N.
FEBS Lett. 315, 114-118, 1993
A;Title: A rat brain cDNA encoding the neurotransmitter transporter with an unusua
A;Reference number: S27043; MUID:93114444; PMID:8093354
A;Accession: S27043
A;Molecule type: mRNA
A;Residues: 1-727 <LIU>
A;Cross-references: GB:S52051; NID:g262842; PIDN:AAB24776.1; PID:g262843
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: transmembrane protein

Query Match 94.6%; Score 3610; DB 2; Length 727;
Best Local Similarity 94.4%; Pred. No. 2.6e-278;
Matches 686; Conservative 13; Mismatches 28; Indels 0; Gaps 0;

```
Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
           |||||||||||||:|||||||||||||||||||||||||||| |||||  |||||||||
Db       1 MPKNSKVTQREHSNEHVTESVADLLALEEPVDYKQSVLNVAGETGGKQKVAEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
           |||||||||||||:|||||||||||||||||||||||||||||:|||||||||||||||||
Db     121 GQRIRRGSIGVWHYVCPRLGGIGFSSCIVCLFVGLYYNVIIGWSVFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
           |||:|||:|||||  ||||||||||||||||||||:|||||||||||||:|||| ||| ||
Db     181 CPVIRNGTVAVVEPECEKSSATTYFWYREALDISNSISESGGLNWKMTVCLLVAWSIVGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
           || ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     241 AVVKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLDKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
           |||||| |||||||| |||||||||||||| |||||||||||||||||| |||||||
Db     301 AATQVFFALGLGFGGVIAFSSYNKQDNNCHFDAALVSFINFFTSVLATLVVFAVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
           |||||||||||||||||:|||||||||||||||||||| || |||||||| |||||
Db     361 IMNEKCVVENAEKILGYLNSNVLSRDLIPPHVNFSHLTTKDYSEMYNVIMTVKEKQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGLGSMIGTMAG 480
```

```
Qy    481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
          |||||||| ||||||||||||||| | |||||||||||||||||||||||||||:||||||
Db    481 ITTPIIDTFKVPKEMFTVGCCVFAFFVGLLFVQRSGNYFVTMFDDYSATLPLTVIVILEN 540

Qy    541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||:|| |||
Db    541 IAVAWIYGTKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVSPPGYSA 600

Qy    601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
          |||||||||||||||| |||||||| ||||||||||:|||||||||||||||||||| ||
Db    601 WIKEEAAERYLYFPNWAMALLITLIAVATLPIPVVFILRHFHLLSDGSNTLSVSYKKGRM 660

Qy    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
          |||||||||||||||||||||||||||||||||||||||||||:| :||| || |||||
Db    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLESSSHPNGRYGSGYLLA 720

Qy    721 STPESEL 727
          |||||||
Db    721 STPESEL 727


RESULT 2
I56506
Na+/Cl(-)-dependent neurotransmitter transporter, brain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C;Accession: I56506
R;el Mestikawy, S.; Giros, B.; Pohl, M.; Hamon, M.; Kingsmore, S.F.; Seldin, M.F.; Caron
J. Neurochem. 62, 445-455, 1994
A;Title: Characterization of an atypical member of the Na+/Cl(-)-dependent transporter f
A;Reference number: I56506; MUID:94125086; PMID:8294906
A;Accession: I56506
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-727 <RES>
A;Cross-references: GB:S68944; NID:g545077; PIDN:AAC60673.1; PID:g545078
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             94.2%;  Score 3595;  DB 2;  Length 727;
  Best Local Similarity   94.1%;  Pred. No. 4.1e-277;
  Matches  684;  Conservative   13;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
          |||||||||||||:|||||||||||||||||||||||||||||| |||||  |||||||||
Db      1 MPKNSKVTQREHSNEHVTESVADLLALEEPVDYKQSVLNVAGETGGKQKVAEEELDAEDR 60

Qy     61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy    121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
          |||||||||||||:|||||||||||||||||||||||||||:|||||||||||||||||||
Db    121 GQRIRRGSIGVWHYVCPRLGGIGFSSCIVCLFVGLYYNVIIGWSVFYFFKSFQYPLPWSE 180

Qy    181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
          |||:|||:||||| |||||||||||||||||:|||||||||||||:|||| ||| ||
Db    181 CPVIRNGTVAVVEPECEKSSATTYFWYREALDISNSISESGGLNWKMTVCLLVAWSIVGM 240

Qy    241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
          || |||||||||||||||| ||||||||||||||||||||||||||||| |||||||||
Db    241 AVVKGIQSSGKVMYFSSLFPCVVLACFLVRGLLLRGAVDGILHMFTPKLDKMLDPQVWRE 300

Qy    301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
          ||||||| |||||||| |||||||||||||| |||||||||||||||||| |||||||
Db    301 AATQVFFALGLGFGGVIAFSSYNKQDNNCHFDAALVSFINFFTSVLATLVVFAVLGFKAN 360

Qy    361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
          ||||||||||||||||||:|||||||||||||||||||| || |||||||| |||||
Db    361 IMNEKCVVENAEKILGYLNSNVLSRDLIPPHVNFSHLTTKDYSEMYNVIMTVKEKQFSAL 420

Qy    421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy    481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
          |||||||| ||||||||| |||| | |||||||||||||||||||||||||:||||||
Db    481 ITTPIIDTFKVPKEMFTVSCCVFAFFVGLLFVQRSGNYFVTMFDDYSATLPLTVIVILEN 540

Qy    541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||:|| |||
Db    541 IAVAWIYGTKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVSPPGYSA 600

Qy    601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
          |||||||||||||||| |||||||| ||||||||||:|||||||||||||||||||| ||
Db    601 WIKEEAAERYLYFPNWAMALLITLIAVATLPIPVVFILRHFHLLSDGSNTLSVSYKKGRM 660

Qy    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
          |||||||||||||||||||||||||||||||||||||||||||:| :||| || |||||
Db    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLESSSHPNGRYGSGYLLA 720

Qy    721 STPESEL 727
          |||||||
Db    721 STPESEL 727


RESULT 3
I52632
sodium-dependent neurotransmitter transporter - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 21-Aug-1998
C;Accession: I52632
R;Uhl, G.R.; Kitayama, S.; Gregor, P.; Nanthakumar, E.; Persico, A.; Shimada, S.
Brain Res. Mol. Brain Res. 16, 353-359, 1992
A;Title: Neurotransmitter transporter family cDNAs in a rat midbrain library: 'orpha
A;Reference number: I52632; MUID:93180651; PMID:1363329
A;Accession: I52632
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-730 <RES>
A;Cross-references: GB:S56141; NID:g298325
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             65.1%;  Score 2483;  DB 2;  Length 730;
  Best Local Similarity   63.9%;  Pred. No. 7.3e-189;
  Matches  472;  Conservative  109;  Mismatches  134;  Indels   24;  Gaps    8;

Qy      2 PKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAED 59
          |||||| :|:   : | ||| ||||: |: |:  |:|  | |  :        |:: ||||
Db      3 PKNSKVVKRD-LDDDVIESVKDLLSNEDSVEDVSKKSELIVDVQE-------EKDTDAED 54

Qy     60 -------RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIP 112
                 |||||||||||||:||||||:|||||||||||||||||:|||:||::||||
Db     55 GSEVDDERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLLVIGIP 114

Qy    113 LFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSF 172
          ||||||:|||||||||||:|| |:||||||:||:|| || ||||||||::||| :||
Db    115 LFFLELSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWTLFYFSQSF 174

Qy    173 QYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
          | |||| :||:|:| |   :| |||||||||:|||||| || ||||||||||| |||
Db    175 QQPLPWDQCPLVKNASHTYIEPECEKSSATTYYWYREALAISSSISESGGLNWKMTGCLL 234

Qy    233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
          | :  :|: ||||||||:|||||||||||| |||:| ||| |::||| |||||||  |
Db    235 AAWVMVCLAMIKGIQSSGKIMYFSSLFPYVVLICFLIRSLLLNGSIDGIRHMFTPKLEMM 294

Qy    293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
          |:|:|||| |||||| |||||||| ||||||:|||||||  ||||||||||||||||||
Db    295 LEPKVWREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVF 354

Qy    353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
          |||||||:|||: :|:| ||  | |  :| |:|| |:| | :| :||  : ::|  |
Db    355 AVLGFKANIVNEKCISQNSEMILKLLKTGNVSWDVIPRHINLSAVTAEDYHVVYDIIQKV 414

Qy    413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
```

```
Db    415 KEEEFAVLHLKACQIEDELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGLG 474

Qy    473 SMIGTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPL 532
Db    475 SMFGTIEGIITPVVDTFKVRKEILTVICCLLAFCIGLMFVQRSGNYFVTMFDDYSATLPL 534

Qy    533 TLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLG 592
Db    535 LIVVILENIAVSFVYGIDKFLEDLTDMLGFAPSKYYYYMWKYISPLMLVTLLIASIVNMG 594

Qy    593 VTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTL 651
Db    595 LSPPGYNAWIKEKASEEFLSYPMWGMVVCFSLMVLAILPVPVVFVIRRCNLIDDSSGNLA 654

Qy    652 SVSYKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNG 711
Db    655 SVTYKRGRVLKEPVNL-DGDDASLIHGKIPSEMSSPNFGKNIYRKQSGSPTLDTA--PNG 711

Qy    712 PYGRGYLLA---STPESEL 727
Db    712 RYGIGYLMADMPDMPESDL 730
```

RESULT 4
I65413
sodium-dependent neurotransmitter transporter - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C;Accession: I65413
R;Uhl, G.R.; Kitayama, S.; Gregor, P.; Nanthakumar, E.; Persico, A.; Shimada, S.
Brain Res. Mol. Brain Res. 16, 353-359, 1992
A;Title: Neurotransmitter transporter family cDNAs in a rat midbrain library: 'orphan tr
A;Reference number: I52632; MUID:93180651; PMID:1363329
A;Accession: I65413
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-728 <RES>
A;Cross-references: GB:S56968; NID:g298323; PIDN:AAB25532.1; PID:g298324
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 64.5%; Score 2463; DB 2; Length 728;
Best Local Similarity 63.6%; Pred. No. 2.8e-187;
Matches 469; Conservative 109; Mismatches 135; Indels 24; Gaps 8;

```
Qy      4 NSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAED-- 59
Db      3 NSKVVKRD-LDDDVIESVKDLLSNEDSVEDVSKKSELIVDVQE-------EKDTDAEDGS 54

Qy     60 -----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLF 114
Db     55 EVDDERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLLVIGIPLF 114

Qy    115 FLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQY 174
Db    115 FLELSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYXVALYYNVIIGWTLFYFSQSFQQ 174

Qy    175 PLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVV 234
Db    175 PLPWDQCPLVKNASHTYIEXECEKSSATTYYWYREALAISSSISESGGLNWKMTGCLLAX 234

Qy    235 WSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLD 294
Db    235 WVMVCLAMIKGIQSSGKIMYFSSLFPYVVLICFLIRSLLLNGSIDGIRHMFTPKLEMMLE 294

Qy    295 PQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVV 354
Db    295 PKVWREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAV 354

Qy    355 LGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKE 414
Db    355 LGFKANIVNEKCISQNSEMILKLLXTGNVSWDVIPRHINLSAVTAEDYHVVYDIIQKVKE 414
```

```
Qy    415 DQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSM 474
Db    415 EEFAVLHLKACQIEDELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGLGSM 474

Qy    475 IGTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTL 534
Db    475 FGTIEGIITPVVDTFKVRKEILTVICCLLAFCIGLMFVQRSGNYFVTMFDDYSATLPLLI 534

Qy    535 IVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVT 594
Db    535 VVILENIAVSFVYGIDKFLEDLTDMLGFAPSKYYYYMWKYISPLMLVTLLIASIVNMGLS 594

Qy    595 PPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSV 653
Db    595 PPGYNAWIKEKASEEFLSYPMWGMVVCFSLMVLAILPVPVVFVIRRCNLIDDSSGNLASV 654

Qy    654 SYKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPY 713
Db    655 TYKRGRVLKEPVNL-DGDDASLIHGKIPSEMSSPNFGKNIYRKQSGSPTLDTA--PNGRY 711

Qy    714 GRGYLLA---STPESEL 727
Db    712 GIGYLMADMPDMPESDL 728
```

RESULT 5
S50998
transporter protein (clone rB21a) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S50998
R;Smith, K.E.; Fried, S.G.; Durkin, M.M.; Gustafson, E.L.; Borden, L.A.; Branchek,
FEBS Lett. 357, 86-92, 1995
A;Title: Molecular cloning of an orphan transporter. A new member of the neurotrans
A;Reference number: S50998; MUID:95094940; PMID:8001687
A;Accession: S50998
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-616 <SMI>
A;Cross-references: GB:S76742; NID:g914027; PIDN:AAB32806.1; PID:g914028
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 34.1%; Score 1302.5; DB 2; Length 616;
Best Local Similarity 43.1%; Pred. No. 2.9e-95;
Matches 265; Conservative 122; Mismatches 191; Indels 37; Gaps 14;

```
Qy     38 LNVAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
Db      3 LAIKRRASRGQRPGPDEKRARDMEKARPQWGNPLQFVFACISYAVGLGNVWRFPYLCQMY 62

Qy     94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
Db     63 GGGSFLVPYLIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFL 122

Qy    154 GLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDI 213
Db    123 SMYYNVINAWGFWYLFHSFQDPLPWSVCPL--NSNRTGYDEECEKASSTQYFWYRKTLNI 180

Qy    214 SDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLL 273
Db    181 SPSIQENGGVQWEPALCLTLAWLMVYLCILRGTESTGKVVYFTALMPYCVLIIYLVRGLT 240

Qy    274 LRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDG 333
Db    241 LHGATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNEPSNDCQKHA 300

Qy    334 ALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVN 393
Db    301 VIVSVINSSTSIFASIVTFSIYGFKATFNYENCL----NKVI-LLLTN--SFDL-----E 348

Qy    394 FSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTH 451
Db    349 DGFLTASNLEEVKDYLASTYPNKYSEVFPHIRNCSLESELNTAVQGTGLAFIVYAEAIKN 408
```

```
Qy    452 FPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLV 507
Db    409 MEVSQLWSVLYFFMLLMLGMGSMLGNTAAILTPLTDSKVISSYLPKEAISGLVCLINCAV 468

Qy    508 GLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRF 567
Db    469 GMVFTMEAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLRAMTG-RPLNW 527

Qy    568 YF-YMWKFVSPLC---MAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWP---MAL 620
Db    528 YWKAMWAFVSPLLIIGLFIFYLSDYILTGTL--QYQAW---DATQGQLVTKDYPPHALAV 582

Qy    621 LITLIVVATLPIPVV 635
Db    583 IGLLVASSTMCIPLV 597


RESULT 6
A48716
glycine transporter GLYT2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 02-Jun-1995
C;Accession: A48716
R;Liu, Q.R.; Lopez-Corcuera, B.; Mandiyan, S.; Nelson, H.; Nelson, N.
J. Biol. Chem. 268, 22802-22808, 1993
A;Title: Cloning and expression of a spinal cord- and brain-specific glycine transporter
A;Reference number: A48716; MUID:94043045; PMID:8226790
A;Accession: A48716
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-799 <LIU>
A;Cross-references: GB:L21672
C;Genetics:
A;Gene: GLYT2
C;Keywords: brain; transmembrane protein

  Query Match             26.9%;  Score 1028;  DB 2;  Length 799;
  Best Local Similarity   32.2%;  Pred. No. 2.5e-73;
  Matches  228;  Conservative  118;  Mismatches  208;  Indels  154;  Gaps   16;

Qy     43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
Db    185 EQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQNGGGAFLIPY 235

Qy    103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
Db    236 LMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVLIAIYYNVIIC 294

Qy    163 WSIFYFFKSFQYPLPWSEC-------------------------PVVR------------ 185
Db    295 YTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVIGDHPKIQIKNSTFCMTAYP 354

Qy    186 NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAV 242
Db    355 NLTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLPFCLFLAWVIVYASL 413

Qy    243 GKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVA 302
Db    414 AKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKLTDATVWKDAA 473

Qy    303 TQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIM 362
Db    474 TQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIFSVIGFMAN-- 531

Qy    363 NEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGL 422
Db    532 ERKVNIENV---------------------------------------------------- 540

Qy    423 DPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGIT 482
Db    541 -----------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLDTMFATIETIV 589

Qy    483 TPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVIL 538
Db    590 TSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAASYALVIIAIF 649

Qy    539 ENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAY 598
Db    650 ELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSFYQW--EPMTY 707

Qy    599 SAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKA 658
Db    708 GSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIKMY--LAPG---------- 744

Qy    659 RMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETS 706
Db    745 RFI---------ERLKLVCSPQPDWGPF-LAQHRGERYKNMIDPLGTS 782


RESULT 7
I58140
glycine transporter - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C;Accession: I58140
R;Borowsky, B.; Mezey, E.; Hoffman, B.J.
Neuron 10, 851-863, 1993
A;Title: Two glycine transporter variants with distinct localization in the CNS and
A;Reference number: I58140; MUID:93264081; PMID:8494645
A;Accession: I58140
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-633 <RES>
A;Cross-references: GB:L13600; NID:g294557; PIDN:AAA73557.1; PID:g294558
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             26.3%;  Score 1004;  DB 2;  Length 633;
  Best Local Similarity   31.4%;  Pred. No. 1.5e-71;
  Matches  224;  Conservative  130;  Mismatches  224;  Indels  136;  Gaps   16;

Qy     42 GEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGA 97
Db      5 GAKGMLNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGA 64

Qy     98 YLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYY 157
Db     65 FMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYY 123

Qy    158 NVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-VVE 193
Db    124 NVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHLFN 183

Qy    194 AECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVM 253
Db    184 YTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGKVV 243

Qy    254 YFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGF 313
Db    244 YFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGCAW 303

Qy    314 GGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEK 373
Db    304 GGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN------------- 350

Qy    374 ILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDK 433
Db    351 ---------------------------------------------HLGVDVSRVAD---- 361

Qy    434 SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPLIDTSKVPK 493
Db    362 --HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGN 417

Qy    494 EMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWI 546
Db    418 EWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIMYI 477
```

```
Qy    547 YGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEA 606
          || : : |:: |||| |  |:   |:|||| :  :   ::||    |  |:
Db    478 YGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN------- 528

Qy    607 AERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDI 664
             :  :| | :|:  |: |  |  :|:  :| |      :||   |       :  :|
Db    529 ---HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSRDW 581

Qy    665 S-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
            | |:   |:     |:  |||       | |        :|   :| |: |
Db    582 GPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 631


RESULT 8
JH0673
glycine transport protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 20-Aug-1999
C;Accession: JH0673; A46204
R;Smith, K.E.; Borden, L.A.; Hartig, P.R.; Branchek, T.; Weinshank, R.L.
Neuron 8, 927-935, 1992
A;Title: Cloning and expression of a glycine transporter reveal colocalization with NMDA
A;Reference number: JH0673; MUID:92265306; PMID:1534013
A;Accession: JH0673
A;Molecule type: mRNA
A;Residues: 1-638 <SMI>
A;Cross-references: GB:M88595; NID:g204433; PIDN:AAA41257.1; PID:g204434
A;Experimental source: brain
R;Guastella, J.; Brecha, N.; Weigmann, C.; Lester, H.A.; Davidson, N.
Proc. Natl. Acad. Sci. U.S.A. 89, 7189-7193, 1992
A;Title: Cloning, expression, and localization of a rat brain high-affinity glycine tran
A;Reference number: A46204; MUID:92357797; PMID:1353889
A;Accession: A46204
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 'M',3,'GK',6,'AKGML',16-638 <GUA>
A;Cross-references: GB:M95413; NID:g204431; PIDN:AAA41256.1; PID:g204432
A;Experimental source: C6 glioma cells and tissues
A;Note: sequence extracted from NCBI backbone (NCBIP:113358)
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: glycoprotein; transmembrane protein
F;41-61/Domain: transmembrane #status predicted <TM1>
F;68-88/Domain: transmembrane #status predicted <TM2>
F;112-132/Domain: transmembrane #status predicted <TM3>
F;222-242/Domain: transmembrane #status predicted <TM4>
F;247-267/Domain: transmembrane #status predicted <TM5>
F;300-320/Domain: transmembrane #status predicted <TM6>
F;329-349/Domain: transmembrane #status predicted <TM7>
F;382-402/Domain: transmembrane #status predicted <TM8>
F;430-450/Domain: transmembrane #status predicted <TM9>
F;466-486/Domain: transmembrane #status predicted <TM10>
F;502-522/Domain: transmembrane #status predicted <TM11>
F;544-564/Domain: transmembrane #status predicted <TM12>
F;169,172,182,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match           26.3%;  Score 1003;  DB 2;  Length 638;
  Best Local Similarity   31.3%;  Pred. No. 1.8e-71;
  Matches  224;  Conservative  132;  Mismatches  224;  Indels  136;  Gaps   16;

Qy     40 VAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
          ||  :  :  ||  |   :|    |  | :::::::|  :|::|||||:|||||||:|||
Db      8 VATSSPEQNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGG 67

Qy     96 GAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGL 155
          ||:: || ::|:  |||||:||: ||   :| :|||  | |   |:|:   :|  ::|:
Db     68 GAFMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGI 126

Qy    156 YYNVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-V 191
          ||||:|  : :||| |  : |||: |                      |   :|::: :
Db    127 YYNVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHL 186

Qy    192 VEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGK 251
```

```
Db    187 FNYTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGK 246

Qy    252 VMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGL 311
          |:||:: ||||||    |||: | ||  ||::  |||  |:|: :|| : |:|:|: ||
Db    247 VVYFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGC 306

Qy    312 GFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENA 371
           :||:|  :|||| |||: |  ::|  |  ||| |  |:| :||| ||
Db    307 AWGGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN----------- 355

Qy    372 EKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDEL 431
                                                          ||:|'   : |
Db    356 ----------------------------------------------HLGVDVSRVAD-- 366

Qy    432 DKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKV 491
              | ||||:|: ||:|  | || ||::|| ||| ||||:    :  : | |:|  :|
Db    367 ----HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EV 420

Qy    492 PKEMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
            |          |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  :::
Db    421 GNEWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIM 480

Qy    545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
          :||| : : |::  |||| |  |:   |:|||| :  :   ::||    |  |:
Db    481 YIYGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN----- 533

Qy    605 EAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMK 662
             :  :| | :|:  |: |  |  :|:  :| |      :||   |       :  :
Db    534 -----HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSR 584

Qy    663 DIS-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
          |    | |:   |:     |:  |||       | |        :|   :| |: |
Db    585 DWGPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 636


RESULT 9
I77912
glycine transporter type 1c - human
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I77912
R;Kim, K.M.; Kingsmore, S.F.; Han, H.; Yang-Feng, T.L.; Godinot, N.; Seldin, M.F.;
Mol. Pharmacol. 45, 608-617, 1994
A;Title: Cloning of the human glycine transporter type 1: molecular and pharmacolo
mes.
A;Reference number: I57956; MUID:94239375; PMID:8183239
A;Accession: I77912
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-638 <RES>
A;Cross-references: GB:S70612; NID:g546770; PIDN:AAB30785.1; PID:g546771
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match           26.1%;  Score 997.5;  DB 2;  Length 638;
  Best Local Similarity   32.2%;  Pred. No. 4.9e-71;
  Matches  221;  Conservative  124;  Mismatches  216;  Indels  125;  Gaps   14;

Qy     40 VAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
          ||  :  :  ||  |    |    |  | :::::::|  :|::|||||:|||||||:|||
Db      8 VAPSSPEQNGAVPSEATKRDQNLKRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGG 67

Qy     96 GAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGL 155
          ||:: || ::||  |||||:||: ||   :| :|||  | |   |:|:   :|  ::|:
Db     68 GAFMFPYFIMLIFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGI 126

Qy    156 YYNVIIGWSIFYFFKSFQYPLPWSEC--PVVRNGSVAVVEA------------------ 194
          ||||:|  : :||| |  : |||: |  |   :|::|
Db    127 YYNVVICIAFYYFFSSMTHVLPWAYCNNPWNTHDCAGVLDASNLTNGSRPAALPSNLSHL 186

Qy    195 ---ECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGK 251
               :::| :  :|    | :|| |   | :   :  || | | :  : : :|::||||
Db    187 LNHSLQRTSPSEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWLVVFLCLIRGVKSSGK 246
```

```
Qy    252 VMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGL 311
Db    247 VVYFTATFPYVVLTILFVRGVTLEGAFDGIMYYLTPQWDKILEAKVWGDAASQIFYSLAC 306

Qy    312 GFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENA 371
Db    307 AWGGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN----------- 355

Qy    372 EKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDEL 431
Db    356 ----------------------------------------------HLGVDVSRVAD-- 366

Qy    432 DKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKV 491
Db    367 ----HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EV 420

Qy    492 PKEMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
Db    421 GNEWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVAIM 480

Qy    545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
Db    481 YIYGHRNYFQDIQMMLGFPPPLFFQICWRFVSPAIIFFILVFTVIQ--YQPITYN----- 533

Qy    605 EAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMK 662
Db    534 -----HYQYPGWAVAIGFLMALSSVLCIPLYAMFRL----CRTDGDTLLQRLKNATKPSR 584

Qy    663 DIS-NLEENDETRFILSKVPSEAPSP 687
Db    585 DWGPALLEHRTGRY----APTIAPSP 606
```

RESULT 10
I57956
glycine transporter type 1b - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I57956
R;Kim, K.M.; Kingsmore, S.F.; Han, H.; Yang-Feng, T.L.; Godinot, N.; Seldin, M.F.; Caron
Mol. Pharmacol. 45, 608-617, 1994
A;Title: Cloning of the human glycine transporter type 1: molecular and pharmacological
mes.
A;Reference number: I57956; MUID:94239375; PMID:8183239
A;Accession: I57956
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-692 <RES>
A;Cross-references: GB:S70609; NID:g546768; PIDN:AAB30784.1; PID:g546769
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 26.1%; Score 995.5; DB 2; Length 692;
Best Local Similarity 32.4%; Pred. No. 7.8e-71;
Matches 219; Conservative 122; Mismatches 210; Indels 125; Gaps 14;

```
Qy     50 AVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVL 105
Db     72 AVPSEATKRDQNLKRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGAFMFPYFIM 131

Qy    106 LIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSI 165
Db    132 LIFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYYNVVICIAF 190

Qy    166 FYFFKSFQYPLPWSEC--PVVRNGSVAVVEA---------------------ECEKSSA 201
Db    191 YYFFSSMTHVLPWAYCNNPWNTHDCAGVLDASNLTNGSRPAALPSNLSHLLNHSLQRTSP 250

Qy    202 TTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPY 261
Db    251 SEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWLVVFLCLIRGVKSSGKVVYFTATFPY 310

Qy    262 VVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSS 321
Db    311 VVLTILFVRGVTLEGAFDGIMYYLTPQWDKILEAKVWGDAASQIFYSLACAWGGLITMAS 370

Qy    322 YNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTN 381
Db    371 YNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN---------------------- 409

Qy    382 VLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLA 441
Db    410 ------------------------------------HLGVDVSRVAD------HGPGLA 426

Qy    442 FIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPKEMF----- 496
Db    427 FVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGNEWILQKKT 484

Qy    497 --TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQ 554
Db    485 YVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVAIMYIYGHRNYFQ 544

Qy    555 ELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFP 614
Db    545 DIQMMLGFPPPLFFQICWRFVSPAIIFFILVFTVIQ--YQPITYN----------HYQYP 592

Qy    615 NWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDIS-NLEEND 671
Db    593 GWAVAIGFLMALSSVLCIPLYAMFRL----CRTDGDTLLQRLKNATKPSRDWGPALLEHR 648

Qy    672 ETRFILSKVPSEAPSP 687
Db    649 TGRY----APTIAPSP 660
```

RESULT 11
S14278
noradrenaline transport protein - human
C;Species: Homo sapiens (man)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S14278
R;Pacholczyk, T.; Blakely, R.D.; Amara, S.G.
Nature 350, 350-354, 1991
A;Title: Expression cloning of a cocaine- and antidepressant-sensitive human noradre
A;Reference number: S14278; MUID:91179515; PMID:2008212
A;Accession: S14278
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-617 <PAC>
A;Cross-references: GB:M65105; NID:g189257; PIDN:AAA59943.1; PID:g189258
C;Genetics:
A;Gene: GDB:SLC6A2; NAT1
A;Cross-references: GDB:127367; OMIM:163970
A;Map position: 16q12.2-16q12.2
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: transmembrane protein

Query Match 25.8%; Score 985.5; DB 2; Length 617;
Best Local Similarity 34.5%; Pred. No. 4.2e-70;
Matches 219; Conservative 110; Mismatches 200; Indels 105; Gaps 14;

```
Qy     28 EEPVDYKQSV-LNVAGEAGGKQ-KAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWR 85
Db     22 EQPLRARKTAELLVVKERNGVQCLLAPRDGDAQPRETWGKKIDFLLSVVGFAVDLANVWR 81

Qy     86 FPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFS 145
Db     82 FPYLCYKNGGGAFLIPYTLFLIIAGMPLFYMELALGQYNREGAATVWK-ICPFFKGVGYA 140

Qy    146 SCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSEC----------PVVRNGSVAVVEA 194
Db    141 VILIALYVGFYYNVIIAWSLYYLFSSFTLNLPWTDCGHTWNSPNCTDPKLLNGSVLGNHT 200

Qy    195 ECEKSSAT--TYFWYREALDISDS--ISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSG 250
Db    201 KYSKYKFTPAAEFYERGVLHLHESSGIHDIGLPQWQLLLCLMVVVIVLYFSLWKGVKTSG 260
```

```
Qy    251 KVMYFSSLFPYVVLACFLVRGLLLRGAVDGI---LHMFTPKLVKMLDPQVWREVATQVFF 307
Db    261 KVVWITATLPYFVLFVLLVHGVTLPGASNGINAYLHI---DFYRLKEATVWIDAATQIFF 317

Qy    308 GLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCV 367
Db    318 SLGAGFGVLIAFASYNKFDNNCYRDALLTSSINCITSFVSGFAIFSILGYMAH--EHKVN 375

Qy    368 VENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLL 427
Db    376 IE---------------------------------------------------------- 377

Qy    428 EDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID 487
Db    378 ----DVATEGAGLVFILYPEAISTLSGSTFWAVVFFVMLLALGLDSSMGGMEAVITGLAD 433

Qy    488 TSKVPK---EMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
Db    434 DFQVLKRHRKLFTFGVTFSTFLLALFCITKGGIYVLTLLDTFAAGTSILFAVLMEAIGVS 493

Qy    545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
Db    494 WFYGVDRFSNDIQQMMGFRPGLYWRLCWKFVSPAFLLFVVVVSIINF--KPLTYDDYI-- 549

Qy    605 EAAERYLYFPNWP--MALLITLIVVATLPIPVVF 636
Db    550 --------FPPWANWVGWGIALSSMVLVPIYVIY 575
```

RESULT 12
S20346
dopamine transport protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S20346
R;Giros, B.; el Mestikawy, S.; Bertrand, L.; Caron, M.G.
FEBS Lett. 295, 149-154, 1991
A;Title: Cloning and functional characterization of a cocaine-sensitive dopamine transpo
A;Reference number: S20346; MUID:92111727; PMID:1765147
A;Accession: S20346
A;Molecule type: mRNA
A;Residues: 1-619 <GIR>
A;Cross-references: GB:S76145; NID:g243202; PIDN:AAB21099.1; PID:g243203
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: transmembrane protein

  Query Match 25.7%; Score 981.5; DB 2; Length 619;
  Best Local Similarity 35.6%; Pred. No. 8.7e-70;
  Matches 213; Conservative 100; Mismatches 186; Indels 99; Gaps 11;

```
Qy     55 LDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLF 114
Db     55 VEAQERETWSKKIDFLLSVIGFAVDLANVWRFPYLCYKNGGGAFLVPYLLFMVIAGMPLF 114

Qy    115 FLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQY 174
Db    115 YMELALGQFNREGAAGVWK-ICPVLKGVGFTVILISFYVGFFYNVIIAWALHYFFSSFTM 173

Qy    175 PLPWSECPVVRNGSVAVVEAECEKSS------------ATTYFWYREALDISDS--ISE 219
Db    174 DLPWIHCNNTWN-SPNCSDAHASNSSDGLGLNDTFGTTPAAEYF-ERGVLHLHQSRGIDD 231

Qy    220 SGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVD 279
Db    232 LGPPRWQLTACLVLVIVLLYFSLWKGVKTSGKVVWITATMPYVVLTALLLRGVTLPGAMD 291

Qy    280 GILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFI 339
Db    292 GIRAYLSVDFYRLCEASVWIDAATQVCFSLGVGFGVLIAFSSYNKFTNNCYRDAIITTSI 351

Qy    340 NFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTT 399
Db    352 NSLTSFSSGFVVFSFLGYMAQ------------------KHNVPIRDV------------ 381

Qy    400 KDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWS 459
Db    382 ---------------------------------ATDGPGLIFIIYPEAIATLPLSSAWA 407

Qy    460 VMFFLMLINLGLGSMIGTMAGITTPIIDTSKV---PKEMFTVGCCVFTFLVGLLFVQRSG 516
Db    408 AVFFLMLLTLGIDSAMGGMESVITGLVDEFQLLHRHRELFTLGIVLATFLLSLFCVTNGG 467

Qy    517 NYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVS 576
Db    468 IYVFTLLDHFAAGTSILFGVLIEAIGVAWFYGVQQFSDDIKQMTGQRPNLYWRLCWKLVS 527

Qy    577 PLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMAL--LITLIVVATLPI 632
Db    528 PCFLLYVVVVSIVTF--RPPHYGAYI----------FPDWANALGWIIATSSMAMVPI 573
```

RESULT 13
I59558
dopamine transporter - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I59558; I59559
R;Shimada, S.; Kitayama, S.; Lin, C.
Science 254, 576-578, 1991
A;Title: Cloning and expression of a cocaine-sensitive dopamine transporter cDNA.
A;Reference number: I59558; MUID:92054539; PMID:1948034
A;Accession: I59558
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-619 <RES>
A;Cross-references: GB:M80570; NID:g310097; PIDN:AAA73143.1; PID:g310098
R;Kilty, J.E.; Lorang, D.; Amara, S.G.
Science 254, 578-579, 1991
A;Title: Cloning and expression of a cocaine-sensitive rat dopamine transporter.
A;Reference number: I59559; MUID:92054540; PMID:1948035
A;Accession: I59559
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-619 <RE2>
A;Cross-references: GB:M80233; NID:g203980; PIDN:AAA41100.1; PID:g203981
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match 25.7%; Score 981.5; DB 2; Length 619;
  Best Local Similarity 35.6%; Pred. No. 8.7e-70;
  Matches 213; Conservative 100; Mismatches 186; Indels 99; Gaps 11;

```
Qy     55 LDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLF 114
Db     55 VEAQERETWSKKIDFLLSVIGFAVDLANVWRFPYLCYKNGGGAFLVPYLLFMVIAGMPLF 114

Qy    115 FLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQY 174
Db    115 YMELALGQFNREGAAGVWK-ICPVLKGVGFTVILISFYVGFFYNVIIAWALHYFFSSFTM 173

Qy    175 PLPWSECPVVRNGSVAVVEAECEKSS------------ATTYFWYREALDISDS--ISE 219
Db    174 DLPWIHCNNTWN-SPNCSDAHASNSSDGLGLNDTFGTTPAAEYF-ERGVLHLHQSRGIDD 231

Qy    220 SGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVD 279
Db    232 LGPPRWQLTACLVLVIVLLYFSLWKGVKTSGKVVWITATMPYVVLTALLLRGVTLPGAMD 291

Qy    280 GILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFI 339
Db    292 GIRAYLSVDFYRLCEASVWIDAATQVCFSLGVGFGVLIAFSSYNKFTNNCYRDAIITTSI 351

Qy    340 NFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTT 399
Db    352 NSLTSFSSGFVVFSFLGYMAQ------------------KHNVPIRDV------------ 381
```

```
Qy    400 KDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWS 459
                                                :  | || || : ||:   | |  |:
Db    382 ----------------------------------ATDGPGLIFIIYPEAIATLPLSSAWA 407

Qy    460 VMFFLMLINLGLGSMIGTMAGITTPIIDTSKV---PKEMFTVGCCVFTFLVGLLFVQRSG 516
           :|||||: ||: | :| |  : | ::|  ::    :|:||:|  : |||: |  |   |
Db    408 AVFFLMLLTLGIDSAMGGMESVITGLVDEFQLLHRHRELFTLGIVLATFLLSLFCVTNGG 467

Qy    517 NYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVS 576
           |  |: | ::|   :   |::| | ||| || ::|  :: :| | ||  ::   || ||
Db    468 IYVFTLLDHFAAGTSILFGVLIEAIGVAWFYGVQQFSDDIKQMTGQRPNLYWRLCWKLVS 527

Qy    577 PLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMAL--LITLIVVATLPI 632
          |  :  :   ||:     || | |:|          ||:|  ||  :|   :| ::||
Db    528 PCFLLYVVVVSIVTF--RPPHYGAYI----------FPDWANALGWIIATSSMAMVPI 573


RESULT 14
S23151
glycine transporter protein GLYT1a - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S23151; B48716
R;Liu, Q.R.; Nelson, H.; Mandiyan, S.; Lopez-Corcuera, B.; Nelson, N.
FEBS Lett. 305, 110-114, 1992
A;Title: Cloning and expression of a glycine transporter from mouse brain.
A;Reference number: S23151; MUID:92316254; PMID:1618338
A;Accession: S23151
A;Molecule type: mRNA
A;Residues: 1-633 <LIU>
A;Cross-references: EMBL:X67056; NID:g51092; PIDN:CAA47440.1; PID:g51093
R;Liu, Q.R.; Lopez-Corcuera, B.; Mandiyan, S.; Nelson, H.; Nelson, N.
J. Biol. Chem. 268, 22802-22808, 1993
A;Title: Cloning and expression of a spinal cord- and brain-specific glycine transporter
A;Reference number: A48716; MUID:94043045; PMID:8226790
A;Accession: B48716
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-62 <LI2>
C;Genetics:
A;Gene: GLYT1
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: alternative splicing; transmembrane protein

  Query Match             25.7%;  Score 981.5;  DB 2;  Length 633;
  Best Local Similarity   31.3%;  Pred. No. 9e-70;
  Matches  217;  Conservative  126;  Mismatches  220;  Indels  131;  Gaps   15;

Qy     42 GEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGA 97
          |  |    ||  |   :|    |  | :::::::| :|::|||||:||||||| :|||||
Db      5 GAKGMLNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGA 64

Qy     98 YLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYY 157
          :: || ::||  |||||||:||: ||   :| :|||  | |   |:|:   :|  ::|:||
Db     65 FMFPYFIMLIFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGIYY 123

Qy    158 NVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-VVE 193
          ||:|  : :||| |  : |||: |                      |   :|::: :
Db    124 NVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPAALSGNLSHLFN 183

Qy    194 AECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVM 253
             :::| :  :|    | :|| |   | :   :  || | | :  : : :|::|||||:
Db    184 YTLQRTSPSEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGKVV 243

Qy    254 YFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGF 313
          ||:: ||||||    |||: | ||  ||::  ||:  | |: :|| : |:|:|: ||  :
Db    244 YFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPQWDKYLEAKVWGDAASQIFYSLGCAW 303

Qy    314 GGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEK 373
          ||:|  :|||| |||: |  ::|  |  | : |  |:| :||| ||
Db    304 GGLITMASYNKFHNNCYRDSVIISITNCATRLYAGFVIFSILGFMAN------------ 350

Qy    374 ILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDK 433
                                                        ||:|   : |
Db    351 ----------------------------------------------HLGVDVSRVAD---- 361

Qy    434 SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVPK 493
             | ||||:|: ||:|  | || ||::|| ||| ||||:     :  : | |:|  :|
Db    362 --HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EVGN 417

Qy    494 EMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWI 546
             |         |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  ::: :|
Db    418 EWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIMYI 477

Qy    547 YGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEA 606
             || : : |::  |||| |  |:   |:|||  :  :   ::||    |  |:
Db    478 YGHRNYFQDIQMMLGFPPPLFFQICWRFVSPAIIFFILIFTVIQ--YRPITYN------- 528

Qy    607 AERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDI 664
             :  :| | : :  |: |  |  :|:  :| |      :||   |    ; :  :|
Db    529 ---HYQYPGWAVRIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSRDW 581

Qy    665 S-NLEENDETRFILSKVPSEAPSP------MPTH 691
             | |:   |:     |:  |||       | |
Db    582 GPALLEHRTGRY----APTTTPSPEDGFEVQPLH 611


RESULT 15
A46061
Na(+)-dependent creatine transporter - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C;Accession: A46061
R;Guimbal, C.; Kilimann, M.W.
J. Biol. Chem. 268, 8418-8421, 1993
A;Title: A Na(+)-dependent creatine transporter in rabbit brain, muscle, heart, and
A;Reference number: A46061; MUID:93231984; PMID:8473283
A;Accession: A46061
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-635 <GUI>
A;Experimental source: brain and muscle
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:129571, NCBIP:129580)
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             25.6%;  Score 975.5;  DB 2;  Length 635;
  Best Local Similarity   33.4%;  Pred. No. 2.7e-69;
  Matches  204;  Conservative  108;  Mismatches  188;  Indels  111;  Gaps   14;

Qy     60 RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELA 119
          |  |  :: :|:: :||:|:||||:|||||| ||||| :|:||::: :: |||:||||::
Db     52 RETWTRQMDFIMSCVGFAVGLGNVWRFPYLCYKNGGGVFLIPYVLIALVGGIPIFFLEIS 111

Qy    120 VGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWS 179
          :|| :: ||| ||: |||   |:|::| ::  :   || ::: |  :|  |||   |||:
Db    112 LGQFMKAGSINVWN-ICPLFKGLGYASMVIVFYCNTYYIMVLAWGFYYLVKSFTTTLPWA 170

Qy    180 EC----------PVVR-----NGSVAVVEAE--CEKSSATTYFWYREALDISDSISESGG 222
          |           : |     |||:| : :   |: |    ||  : | :|  :   |
Db    171 TCGHTWNTPDCVEIFRHEDCANGSLANLTCDQLAERRSPVIEFWENKVLRLSGGLEVPGA 230

Qy    223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
          |||::|||||  | :    | ||:::|:||::||:: ||||||   ||||:|| ||:|||:
Db    231 LNWEVTLCLLACWVLVYFCVWKGVKSTGKIVYFTATFPYVVLVVLLVRGVLLPGALDGII 290

Qy    283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
          :   |   |:  |||| :  ||:||   :| | :    |||: :|||: |  ::: ||
Db    291 YYLKPDWSKLRSPQVWIDAGTQIFFSYAIGLGALTALGSYNRFNNNCYKDAIILALINSG 350

Qy    343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDY 402
          ||  |  ||| :||| |                                        |:
Db    351 TSFFAGFVVFSILGFMA----------------------------------------TEQG 371

Qy    403 MEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMF 462
          : :  |                          :  | ||||||:  |:|  | :| |: :|
```

```
Db    372 VHISKV------------------------AESGPGLAFIAYPRAVTLMPVAPLWAALF 406

Qy    463 FLMLINLGLGSMIGTMAGITTPIIDT------SKVPKEMFTVGCCVFTFLVGLLFVQRSG 516
          | ||: ||| |    : |  | ::|        :  :|:    ||   |:: |  |   |
Db    407 FFMLLLLGLDSQFVGVEGFITGLLDLLPASYYFRFQREISVALCCALCFVIDLSMVTDGG 466

Qy    517 NYFVTMFDDYSATLPLTLI--VILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKF 574
           |   :|| |||:   ||:     | :||||:||  :|| ::  |:|:||  :  : | |
Db    467 MYVFQLFDYYSAS-GTTLLWQAFWECVAVAWVYGADRFMDDIACMIGYRPCPWMKWCWSF 525

Qy    575 VSPL-CMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIP 633
           :|| || :     :      |  |:        : |:| | |  |:     : : | :|
Db    526 FTPLVCMGIFIFNIVY---YKPLVYN--------KTYVY-PWWGEAMGWAFALSSMLCVP 573

Qy    634 VVFVLRHFHLL 644
          :       |||
Db    574 L-------HLL 577
```

Search completed: April 21, 2003, 13:58:38
Job time : 51 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 13:57:53 ; Search time 22 Seconds
(without alignments)
2498.695 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 3817
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3817	100.0	727	10	US-09-923-444A-2	Sequence 2, Appli
2	2522.5	66.1	730	9	US-10-156-239-5	Sequence 5, Appli
3	2522.5	66.1	730	10	US-09-741-149-2	Sequence 2, Appli
4	2522.5	66.1	730	10	US-09-795-693-5	Sequence 5, Appli
5	2494	65.3	729	10	US-09-741-149-4	Sequence 4, Appli
6	1468.5	38.5	676	9	US-09-815-923-12	Sequence 12, Appl
7	1433	37.5	624	9	US-10-156-239-24	Sequence 24, Appl
8	1433	37.5	624	10	US-09-795-693-24	Sequence 24, Appl
9	1325	34.7	599	10	US-09-861-846-4	Sequence 4, Appli
10	1275	33.4	610	10	US-09-861-846-2	Sequence 2, Appli
11	1031.5	27.0	797	10	US-09-795-232-2	Sequence 2, Appli
12	937.5	24.6	630	10	US-09-843-598-10	Sequence 10, Appl
13	926.5	24.3	587	9	US-09-815-923-4	Sequence 4, Appli
14	924	24.2	622	10	US-09-843-598-11	Sequence 11, Appl
15	914	23.9	597	9	US-09-815-923-14	Sequence 14, Appl
16	892.5	23.4	727	9	US-09-815-923-10	Sequence 10, Appl
17	872	22.8	671	10	US-09-843-598-5	Sequence 5, Appli
18	868	22.7	671	10	US-09-843-598-7	Sequence 7, Appli
19	813.5	21.3	437	9	US-10-216-441-4	Sequence 4, Appli
20	813.5	21.3	437	10	US-09-818-656A-4	Sequence 4, Appli
21	813.5	21.3	459	9	US-10-216-441-2	Sequence 2, Appli
22	813.5	21.3	459	10	US-09-818-656A-2	Sequence 2, Appli
23	763.5	20.0	556	9	US-09-815-923-6	Sequence 6, Appli
24	674	17.7	421	10	US-09-843-598-9	Sequence 9, Appli
25	506.5	13.3	579	9	US-09-738-626-4648	Sequence 4648, Ap
26	345.5	9.1	224	10	US-09-843-598-6	Sequence 6, Appli
27	193.5	5.1	196	9	US-09-989-442-87	Sequence 87, Appl
28	141	3.7	52	10	US-09-864-761-43363	Sequence 43363, A
29	126	3.3	151	9	US-09-989-442-101	Sequence 101, App
30	122	3.2	672	9	US-09-928-530-5	Sequence 5, Appli
31	122	3.2	672	9	US-10-162-012-30	Sequence 30, Appl
32	115	3.0	1146	10	US-09-824-734-2	Sequence 2, Appli
33	114	3.0	1503	9	US-10-007-706-1	Sequence 1, Appli
34	111	2.9	49	10	US-09-864-761-48618	Sequence 48618, A
35	108.5	2.8	84	10	US-09-864-761-39755	Sequence 39755, A
36	108.5	2.8	1085	9	US-10-274-990-4	Sequence 4, Appli
37	108.5	2.8	1085	10	US-09-734-674-4	Sequence 4, Appli
38	106	2.8	563	10	US-09-915-181A-6	Sequence 6, Appli
39	103.5	2.7	1032	10	US-09-728-137-8	Sequence 8, Appli
40	103.5	2.7	1044	10	US-09-728-137-2	Sequence 2, Appli
41	102.5	2.7	714	10	US-09-728-137-4	Sequence 4, Appli
42	102	2.7	497	10	US-09-815-242-12626	Sequence 12626, A
43	100	2.6	469	10	US-09-815-242-5281	Sequence 5281, Ap
44	99.5	2.6	539	9	US-10-051-902-26	Sequence 26, Appl
45	99.5	2.6	539	9	US-10-051-909-26	Sequence 26, Appl

ALIGNMENTS

```
RESULT 1
US-09-923-444A-2
; Sequence 2, Application US/09923444A
; Patent No. US20020015980A1
;  GENERAL INFORMATION:
;       APPLICANT: LI, YI
;                  FLEISCHMANN, ROBERT
;       TITLE OF INVENTION: NEUROTRANSMITTER TRANSPORTER
;       NUMBER OF SEQUENCES: 6
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: MD
;            COUNTRY: US
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/923,444A
;            FILING DATE: 08-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/062,815
;            FILING DATE: 199-12-20
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Michele M. Wales
;            REGISTRATION NUMBER: 43,975
;            REFERENCE/DOCKET NUMBER: PF116
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (301) 309-8504
;            TELEFAX: (301) 309-8439
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 727 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
```

US-09-923-444A-2

Query Match 100.0%; Score 3817; DB 10; Length 727;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
Db       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
Db     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
Db     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
Db     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
Db     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
Db     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
Db     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
Db     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600

Qy     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
Db     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660

Qy     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
Db     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720

Qy     721 STPESEL 727
Db     721 STPESEL 727
```

RESULT 2
US-10-156-239-5
; Sequence 5, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Trans
; TITLE OF INVENTION: ATPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A Hu
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-156-239-5

Query Match 66.1%; Score 2522.5; DB 9; Length 730;
Best Local Similarity 64.4%; Pred. No. 3.7e-225;
Matches 473; Conservative 114; Mismatches 135; Indels 13; Gaps 8;

```
Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
Db       1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy      59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
Db      59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy     118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
Db     119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy     178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237
Db     179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy     238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297
Db     239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy     298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357
Db     299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy     358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417
Db     359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy     418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
Db     419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy     478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
Db     479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy     538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
Db     539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy     598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
Db     599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy     657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
```

```
Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727
          ||:|      |||:|
Db    716 YLMADIMPDMPESDL 730
```

RESULT 3
US-09-741-149-2
; Sequence 2, Application US/09741149
; Patent No. US20020031800A1
; GENERAL INFORMATION:
; APPLICANT: LI, Zhenya et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000780
; CURRENT APPLICATION NUMBER: US/09/741,149
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Human
US-09-741-149-2

Query Match 66.1%; Score 2522.5; DB 10; Length 730;
Best Local Similarity 64.4%; Pred. No. 3.7e-225;
Matches 473; Conservative 114; Mismatches 135; Indels 13; Gaps 8;

```
Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
          ||||||| :||   : ||||| ||| |:  |  :| | | | |:   |   |||  : |
Db      1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy     59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
          | |||||||||||||:||||||||:|||||||||||||||||:|||:||::|||||||||
Db     59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
          |:|||||||||||||:|| |:||||||:||:|| || ||||||||||:||| :||| |||
Db    119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy    178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237
          | :||:|:| |   || |||:||||||:||||||:|| |||||||||||||:|||  | :
Db    179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy    238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297
            :|: ||||||||::|||||||||| |||:|  || |::||| |||||||  ||:|:|
Db    239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy    298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357
          ||| |||||| |||||||| ||||||:|||||||  ||||||||||||||||| ||||
Db    299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy    358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417
          |||::|||||: :|:| |: :|    :|:|:|| |:| | :| :||  : ::|  |||::|
Db    359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy    418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
            || |: | :|:||:|:|||||||||||||||||| ||||||||||:|||||| ||
Db    419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy    478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
          : || |||:|| || ||: || ||:  | :||:|||||||||||||||||| ::||
Db    479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy    538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
          |||||| ::||  |||::| :|||| | |:|:||||::||| :  |  ||:: :|::||
Db    539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy    598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
          |:|||:::|:| :| :| | : : ::|:| | ||:||||::| |:|: | | |  ||:||
```

```
Db    599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy    657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
          : |::|:  || | |:|  |  |:||| |||      |      :  |:|:  ||| || |
Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727
          ||:|      |||:|
Db    716 YLMADIMPDMPESDL 730
```

RESULT 4
US-09-795-693-5
; Sequence 5, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; TITLE OF INVENTION: 32613, No. US20020068710A1el Human Transporters
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-693-5

Query Match 66.1%; Score 2522.5; DB 10; Length 730;
Best Local Similarity 64.4%; Pred. No. 3.7e-225;
Matches 473; Conservative 114; Mismatches 135; Indels 13; Gaps 8;

```
Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
          ||||||| :||   : ||||| ||| |:  |  :| | | | |:   |   |||  : |
Db      1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy     59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
          | |||||||||||||:||||||||:|||||||||||||||||:|||:||::|||||||||
Db     59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
          |:|||||||||||||:|| |:||||||:||:|| || ||||||||||:||| :||| |||
Db    119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy    178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237
          | :||:|:| |   || |||:||||||:||||||:|| |||||||||||||:|||  | :
Db    179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy    238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297
            :|: ||||||||::|||||||||| |||:|  || |::||| |||||||  ||:|:|
Db    239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy    298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357
          ||| |||||| |||||||| ||||||:|||||||  ||||||||||||||||| ||||
Db    299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy    358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417
          |||::|||||: :|:| |: :|    :|:|:|| |:| | :| :||  : ::|  |||::|
Db    359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy    418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
            || |: | :|:||:|:|||||||||||||||||| ||||||||||:|||||| ||
Db    419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy    478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
          : || |||:|| || ||: || ||:  | :||:|||||||||||||||||| ::||
Db    479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy    538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
```

```
Db    539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy    598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
          |:|||:::|:| :| :| | : : ::|:| | ||:||||::| |:|: | | |  ||:||
Db    599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy    657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
          : |::|:  || | |:|  |  |:||| |||      |     :  |:|:  ||| || |
Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727
          ||:|      |||:|
Db    716 YLMADIMPDMPESDL 730


RESULT 5
US-09-741-149-4
; Sequence 4, Application US/09741149
; Patent No. US20020031800A1
; GENERAL INFORMATION:
;  APPLICANT: LI, Zhenya et al
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000780
;  CURRENT APPLICATION NUMBER: US/09/741,149
;  CURRENT FILING DATE:  2000-12-21
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 729
;   TYPE: PRT
;   ORGANISM: Bos taurus
US-09-741-149-4

  Query Match             65.3%;  Score 2494;  DB 10;  Length 729;
  Best Local Similarity   64.0%;  Pred. No. 1.6e-222;
  Matches  471;  Conservative  115;  Mismatches  134;  Indels   16;  Gaps    9;

Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQS--VLNVAGEAGGKQKAVEEELD 56
          ||||||| :||   | | ||| |||: |:  |  :|:|  :::|  |   |   | |  :
Db      1 MPKNSKVVKRELDDE-VIESVKDLLSNEDSADDAFKKSELIVDVPEE---KDTDVVERSE 56

Qy     57 AED-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFF 115
          :| ||||||||||||||:||||||||:|||||||||||||||||:|||:||::|||||||
Db     57 VKDARPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLLVIGIPLFF 116

Qy    116 LELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYP 175
          |||:|||||||||||||:|| |:||||||:||:|| || ||||||||||:||| :||| |
Db    117 LELSVGQRIRRGSIGVWNYISPQLGGIGFASCVVCFFVALYYNVIIGWSLFYFSQSFQQP 176

Qy    176 LPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVW 235
          ||| :||:|:| |   || ||||||||||:||||||:|| |||||||||||||:||| |
Db    177 LPWDQCPLVKNASHTFVEPECEKSSATTYYWYREALNISTSISESGGLNWKMTICLLAAW 236

Qy    236 SIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDP 295
          :  :|: |||||||:||||||||||| |||:| ||| |:|||| |||||:|  ||:|
Db    237 VVVCLAMIKGIQSSGKIMYFSSLFPYVVLICFLIRALLLNGSVDGIRHMFTPELEIMLEP 296

Qy    296 QVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVL 355
          :|||| | |||| |||||||| ||||||:|||||||  ||||||||||:|||||| ||
Db    297 KVWREAAAQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSILATLVVFAVL 356

Qy    356 GFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKED 415
          |||||::||||: ||:| |:  :    :|:|:|:|| |:||| :| :||  : ::|  |||:
Db    357 GFKANVINEKCIAENSEMIIKLVKMGNISQDIIPHHINFSAITAEDYDLIYDIIQKVKEE 416

Qy    416 QFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMI 475
          :| || |: | :||||:|:||||||||||||||||||| ||||||||||||:|||||||
Db    417 EFPALHLNACQIEDELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMF 476

Qy    476 GTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLI 535
```

```
          ||: || ||::|| || ||: || ||:  | :|||:||||||||||||||||||||| ::
Db    477 GTIEGIITPVVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIV 536

Qy    536 VILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTP 595
          ||||||||:::|| |||::| :|||| | |:|:||||::||| :  |  |||: :|::|
Db    537 VILENIAVSFVYGIDKFMEDLKDMLGFTPNRYYYYMWKYISPLMLLSLLIASIVNMGLSP 596

Qy    596 PAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVS 654
          | |:||::::|:|:::| :| | | : |:|:|:| |||||||::|  :|: | | |  ||:
Db    597 PGYNAWMEDKASEKFLSYPTWGMVICISLMVLAILPIPVVFIIRRCNLIDDSSGNLASVT 656

Qy    655 YKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYG 714
          ||: |::|:  || | |:   |  |: ||  ||      |     :  |:|:  ||| ||
Db    657 YKRGRVLKEPVNL-EGDDASLIHGKISSEMSSPNFGKNIYRKQSGSPTLDTA--PNGRYG 713

Qy    715 RGYLLA---STPESEL 727
           |||:|     |||:|
Db    714 IGYLMADMPDMPESDL 729


RESULT 6
US-09-815-923-12
; Sequence 12, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
;  APPLICANT: Gill, Sarjeet S.
;  APPLICANT:  Ross, Linda S.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644.
;  TITLE OF INVENTION:  Target Sites for Insecticides
;  FILE REFERENCE: 02307O-093800US
;  CURRENT APPLICATION NUMBER: US/09/815,923
;  CURRENT FILING DATE:  2001-03-23
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 676
;   TYPE: PRT
;   ORGANISM: Manduca sexta
;   FEATURE:
;   OTHER INFORMATION: orphan transporter
US-09-815-923-12

  Query Match             38.5%;  Score 1468.5;  DB 9;  Length 676;
  Best Local Similarity   44.8%;  Pred. No. 1.9e-127;
  Matches  291;  Conservative  113;  Mismatches  196;  Indels   49;  Gaps   11;

Qy     46 GKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVL 105
          | : ::: :   ||| ||: ||: |: ||:|||||||||||||||:|||||:|:|:|::
Db     38 GSRGSLQRDAPEEDRAAWSGKLQFFLSIIGYSVGLGNIWRFPYLCQQNGGGAFLIPFLIM 97

Qy    106 LIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSI 165
          |:: ||||  :|:|:||::| ||:|||: | | ||||| |||:| ||| |||||||| |
Db     98 LVLEGIPLVLIEMAIGQKMRLGSLGVWNTIHPWLGGIGISSCVVTLFVALYYNVIITWVF 157

Qy    166 FYFFKSFQY---PLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGG 222
          || | | :     |||: ||   ||:    |||| |:||| |||||||||| | || | |
Db    158 FYLFNSIRLTADQLPWAHCP-YDNGT---AEAECNKASATVYFWYREALDASPSIDEPGV 213

Qy    223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
            | : | ||: | |    | |||||||||:||:||||| ||  | |||: | |: ||||
Db    214 PRWWIVLYLLLAWIIVFFIVMKGIQSSGKVVYFTSLFPYAVLTIFFVRGITLPGSSDGIL 273

Qy    283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
          ||: ||| |:||| || : |||||:  || || :| | |||  :|||  |  |||  |
Db    274 HMYKPKLEKLLDPTVWLDAATQVFYSFGLAFGSLIAFGSYNPPNNNCVRDVLLVSVCNAL 333

Qy    343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHV-NFSHLTTKD 401
          |:: |::|:| :|||||  | | |:|:   |:|       .          |:  |:  :| |
Db    334 TAIYASVVIFSILGFKAYTMVENCIVKEI-KVLAL------------HHIGGFTLNSTAD 380

Qy    402 YMEMDNVIMTVKEDQF--------SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFP 453
          |            :||        :|| |  | :  :|::: :|||||| ||:|:
```

```
Db    381 YY----------REQFPRLNGTAIAALNLTGCTMSRQLEEAAEGTGLAFIVFTQAILKLT 430

Qy    454 TSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDT---SKVPKEMFTVGCCVFTFLVGLL 510
          :||||::|||||::|||||| || | |:   | |     :: | : |   | | | |||:
Db    431 PAPFWSIIFFLMLLSLGLGSQIGIMEGMLCTIFDIDFFKRLSKPVITGVVCTFCFFVGLI 490

Qy    511 FVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFY 570
          |   :| |:: ||| :: |: | :: :|: ||| :||| :||  :: || |:||  ::
Db    491 FTTGAGEYWLKMFDSFAGTIGLVVVALLKMIAVIYIYGHEKFTNDIYEMTGYRPGIYWQV 550

Qy    571 MWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATL 630
          |::| |  :  :  :|:: : : || | ||  :|       :| | : : : :|:   |
Db    551 TWRYVGPAIVTCILLSSLVFMLINPPMYGAWNADEGRVIKTPYPTWVLVIAVLMILAGVL 610

Qy    631 PIPVVFVLRHFHLLS------DGS-NTLSVSYKKARMMKDISNLEENDE 672
          ||||| :|| |  |:       ||   :  :     || |  |:| ::|
Db    611 PIPVVLLLRRFQCLAFDVDIHQGSIRRIETTVSTKEMMSDQDNVESSEE 659


RESULT 7
US-10-156-239-24
; Sequence 24, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Trans
;  TITLE OF INVENTION:  ATPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A Hu
;  TITLE OF INVENTION:  Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
;  FILE REFERENCE: 35800/247645
;  CURRENT APPLICATION NUMBER: US/10/156,239
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: 09/795,693
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/185,906
;  PRIOR FILING DATE: 2000-02-29
;  PRIOR APPLICATION NUMBER: 09/809,557
;  PRIOR FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/192,018
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 09/808,568
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/191,790
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 09/808,767
;  PRIOR FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/191,781
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 624
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-24

  Query Match             37.5%;  Score 1433;  DB 9;  Length 624;
  Best Local Similarity   48.0%;  Pred. No. 3.3e-124;
  Matches  306;  Conservative   95;  Mismatches  182;  Indels   54;  Gaps   17;

Qy     60 RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELA 119
          |  |: || ::|: :||:|||||:||||||| ||||||:|:|||: ||: ||||||||||
Db      1 RETWSGKLDFVLSVVGFAVGLGNVWRFPYLCYKNGGGAFLIPYLIFLIVAGIPLFFLELA 60

Qy    120 VGQRIRRGSIGVWHY--------ICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKS 171
          :||  | ||| ||          |||   |||::| ::  ::|:|||||| |:::| | |
Db     61 LGQYTREGSITVWRKKILDKGKGICPLFKGIGYASIVIAFYIGIYYNVIIAWALYYLFSS 120

Qy    172 FQYPLPWSEC---------------PVVRNGSVAVVEAE------CEKSSATTYFWYREA 210
          |   |||: |                   |||:| : ::       |::|    || |
Db    121 FTTELPWATCNNSWNTPNCVEEREAENSTNGSLAALSSKNLTDYTLERTSPVEEFWERGV 180
```

```
Qy    211 LDISDS--ISESGGLNWKMTLCLLVVWSIGGMAVGKGIQS-SGKVMYFSSLFPYVVLACF 267
          | :|:|  | : | | |::|||||: | :    : ||::| ||||:||:: ||||||
Db    181 LKLSESSGIEDLGELRWELTLCLLLAWIVVYFCLWKGVKSGSGKVVYFTATFPYVVLIVL 240

Qy    268 LVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDN 327
          |:||: | || |||    ||   |:|||||| : |||:|| ||:||| :|  :|||  |
Db    241 LIRGVTLPGAADGIKFYLTPDFSKLLDPQVWIDAATQIFFSLGIGFGVLIALASYNKFHN 300

Qy    328 NCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDL 387
          ||: |  :|||||  || ||  |:| :||| |||: |: | || |||   |  :||||||
Db    301 NCYRDAIIVSFINSITSFLAGFVIFSILGFMANIVQEQGVPEN-EKI---LLLSVLSRDL 356

Qy    388 IPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTE 447
          | |||| | ||  ||   | ||  | | :|  |||    |||||||   | ||||||: |
Db    357 I-PHVNLSALTA-DYSVYD-VISEVAESEF-VLGL--ACLEDELDKVQAGPGLAFIAYPE 410

Qy    448 AMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID-----TSKVPKEMFTVGCCV 502
          |:|  | ||||:|:|||||: ||| |  | : || | ::|       || :|:| :  ||
Db    411 AVTMLPLSPFWAVLFFLMLLTLGLDSQFGGVEGIITALVDEFPILLRKVRRELFILLVCV 470

Qy    503 FTFLVGLLFVQRSGNYFVTMFDDYSAT-LPLTLIVILENIAVAWIYGPKKFMQELTEMLG 561
           :||:||  |   | |  |:|| |:|:   |  :|  | |||||:|| :|  ::||||||
Db    471 ISFLLGLFMVTEGGIYVFTLFDYYAASGFSLLFVVFFECIAVAWVYGIDRFYDDITEMLG 530

Qy    562 FRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALL 621
          |||  ::   |||||||  :  |   ||:| |: |  |: |||  || |:| ||| ||
Db    531 FRPGLYWKLCWKFVSPLILLFLFIFSIVQYGLKPLTYNNWIKE--AEDYVY-PNWANALG 587

Qy    622 ITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKA 658
            | : : | :|:  :   : |||   ::|    :||
Db    588 WLLALSSMLCVPLYII---YKLLSTEGDSLLERLQKA 621


RESULT 8
US-09-795-693-24
; Sequence 24, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION:  20685, 579, 17114, 23821, 33894, and
;  TITLE OF INVENTION:  32613, No. US20020068710A1el Human Transporters
;  FILE REFERENCE:  35800/209292
;  CURRENT APPLICATION NUMBER: US/09/795,693
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  60/185,906
;  PRIOR FILING DATE:  2000-02-29
;  NUMBER OF SEQ ID NOS:  42
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 624
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-24

  Query Match             37.5%;  Score 1433;  DB 10;  Length 624;
  Best Local Similarity   48.0%;  Pred. No. 3.3e-124;
  Matches  306;  Conservative   95;  Mismatches  182;  Indels   54;  Gaps   17;

Qy     60 RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELA 119
          |  |: || ::|: :||:|||||:||||||| ||||||:|:|||: ||: ||||||||||
Db      1 RETWSGKLDFVLSVVGFAVGLGNVWRFPYLCYKNGGGAFLIPYLIFLIVAGIPLFFLELA 60

Qy    120 VGQRIRRGSIGVWHY--------ICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKS 171
          :||  | ||| ||          |||   |||::| ::  ::|:|||||| |:::| | |
Db     61 LGQYTREGSITVWRKKILDKGKGICPLFKGIGYASIVIAFYIGIYYNVIIAWALYYLFSS 120

Qy    172 FQYPLPWSEC---------------PVVRNGSVAVVEAE------CEKSSATTYFWYREA 210
          |   |||: |                   |||:| : ::       |::|    || |
Db    121 FTTELPWATCNNSWNTPNCVEEREAENSTNGSLAALSSKNLTDYTLERTSPVEEFWERGV 180
```

```
Qy    211 LDISDS--ISESGGLNWKMTLCLLVVWSIGGMAVGKGIQS-SGKVMYFSSLFPYVVLACF 267
Db    181 LKLSESSGIEDLGELRWELTLCLLLAWIVVYFCLWKGVKSGSGKVVYFTATFPYVVLIVL 240

Qy    268 LVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDN 327
Db    241 LIRGVTLPGAADGIKFYLTPDFSKLLDPQVWIDAATQIFFSLGIGFGVLIALASYNKFHN 300

Qy    328 NCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDL 387
Db    301 NCYRDAIIVSFINSITSFLAGFVIFSILGFMANIVQEQGVPEN-EKI---LLLSVLSRDL 356

Qy    388 IPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTE 447
Db    357 I-PHVNLSALTA-DYSVYD-VISEVAESEF-VLGL--ACLEDELDKVQAGPGLAFIAYPE 410

Qy    448 AMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID-----TSKVPKEMFTVGCCV 502
Db    411 AVTMLPLSPFWAVLFFLMLLTLGLDSQFGGVEGIITALVDEFPILLRKVRRELFILLVCV 470

Qy    503 FTFLVGLLFVQRSGNYFVTMFDDYSAT-LPLTLIVILENIAVAWIYGPKKFMQELTEMLG 561
Db    471 ISFLLGLFMVTEGGIYVFTLFDYYAASGFSLLFVVFFECIAVAWVYGIDRFYDDITEMLG 530

Qy    562 FRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALL 621
Db    531 FRPGLYWKLCWKFVSPLILLFLFIFSIVQYGLKPLTYNNWIKE--AEDYVY-PNWANALG 587

Qy    622 ITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKA 658
Db    588 WLLALSSMLCVPLYII---YKLLSTEGDSLLERLQKA 621
```

RESULT 9
US-09-861-846-4
; Sequence 4, Application US/09861846
; Patent No. US20020110852A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001065
; CURRENT APPLICATION NUMBER: US/09/861,846
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/752,821
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-861-846-4

Query Match 34.7%; Score 1325; DB 10; Length 599;
Best Local Similarity 43.1%; Pred. No. 3.3e-114;
Matches 262; Conservative 115; Mismatches 191; Indels 40; Gaps 11;

```
Qy     43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
Db      3 QASGMDPLVDIE---DERPKWDNKLQYLLSCIGFAVGLGNIWRFPYLCQTHGGGAFLIPY 59

Qy    103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
Db     60 FIALVFEGIPLFYIELAIGQRLRRGSIGVWKTISPYLGGVGLGCFSVSFLVSLYYNTVLL 119

Qy    163 WSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGG 222
Db    120 WVLWFFLNSFQHPLPWSTCPLDLNRTGFV--QECQSSGTVSYFWYRQTLNITSDISNTGT 177

Qy    223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
```

```
Db    178 IQWKLFLCLVACWSTVYLCVIRGIESTGKVIYFTALFPYLVLTIFLIRGLTLPGATEGLI 237

Qy    283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
Db    238 YLFTPNMKTLQNPRVWLDAATQIFFSLSLAFGGHIAFASYNPPRNNCEKDAVIIALVNSM 297

Qy    343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDY 402
Db    298 TSLYASIAIFSVMGFKASNDYGRCLDRN---ILSLIN----EFDL--PELSIS----RD- 343

Qy    403 MEMDNVIMTVKEDQ---FSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWS 459
Db    344 -EYPSVLMYLNATQTARVAQLPLKTCHLEDFLDKSASGPGLAFIVFTEAVLHMPGASVWS 402

Qy    460 VMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRS 515
Db    403 VLFFGMLFTLGLSSMFGNMEGVITPLLDMGILPKGIPKEVMTGVICFACFLSAICFTLQS 462

Qy    516 GNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFV 575
Db    463 GGYWLEIFDSFAASLNLIIFAFMEVVGVIHIYGMKRFCDDIEWMTGRRPGLYWQVTWRVV 522

Qy    576 SPLCMAVLTTASIIQLGVTPPAYSAW--------IKEEAAERYLYFPNWPMALLITLIVV 627
Db    523 SPMLLFGIFLSYIVLLIQTPPSYKAWNPQYEHFPSREEK-----FYPGWVQVTCVLLSFL 577

Qy    628 ATLPIPVV 635
Db    578 PSLWVPGV 585
```

RESULT 10
US-09-861-846-2
; Sequence 2, Application US/09861846
; Patent No. US20020110852A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001065
; CURRENT APPLICATION NUMBER: US/09/861,846
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/752,821
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Human
US-09-861-846-2

Query Match 33.4%; Score 1275; DB 10; Length 610;
Best Local Similarity 42.0%; Pred. No. 1.5e-109;
Matches 250; Conservative 111; Mismatches 190; Indels 44; Gaps 10;

```
Qy     58 EDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
Db     16 DERPKWDNKAQYLLSCIGFAVGLGNIWRFPYLCQTYGGGAFLIPYVIALVFEGIPIFHVE 75

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
Db     76 LAIGQRLRKGSVGVWTAISPYLSGVGLGCVTLSFLISLYYNTIVAWVLWYLLNSFQHPLP 135

Qy    178 WSECP--VVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVW 235
Db    136 WSSCPPDLNRTGFV----EECQGSSAVSYFWYRQTLNITADINDSGSIQWWLLICLAASW 191

Qy    236 SIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDP 295
Db    192 AVVYMCVIRGIETTGKVIYFTALFPYLVLTIFLIRGLTLPGATKGLIYLFTPNMHILQNP 251
```

```
Qy    296 QVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVL 355
Db    252 RVWLDAATQIFFSLSLAFGGHIAFASYNSPRNDCQKDAVVIALVNRMTSLYASIAVFSVL 311

Qy    356 GFKANIMNEKCVVENAEKILGYLN-----TNVLSRDLIPP---HVNFSHLTTKDYMEMDN 407
Db    312 GFKATNDYEHCLDRN---ILSLINDFDFPEQSISRDDYPAVLMHLN-------------- 354

Qy    408 VIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLI 467
Db    355 ---ATWPKRVAQLPLKACLLEDFLDKSASGPGLAFVVFTETDLHMPGAPVWAMLFFGMLF 411

Qy    468 NLGLGSMIGTMAGITTPIIDTSK----VPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMF 523
Db    412 TLGLSTMFGTVEAVITPLLDVGVLPRWVPKEALTGLVCLVCFLSATCFTLQSGNYWLEIF 471

Qy    524 DDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVL 583
Db    472 DNFAASPNLLLMLAFLEVVGVVYVYGMKRFCDDIAWMTGRRPSPYWRLTWRVVSPLLLTIF 531

Qy    584 TTASIIQLGVTPPAYSAWIKE----EAAERYLYFPNWPMALLITLIVVATLPIPV 634
Db    532 -VAYIILLFWKPLRYKAWNPKYELFPSRQEKLY-PGWARAACVLLSLLPVLWVPV 584


RESULT 11
US-09-795-232-2
; Sequence 2, Application US/09795232
; Patent No. US20010012627A1
; GENERAL INFORMATION:
;  APPLICANT:  Anthony M. Brown
;  APPLICANT:  Conrad Gerald Chapman
;  APPLICANT:  Israel Simon Gloger
;  APPLICANT:  Joanne Rachel Evans
;  APPLICANT:  William Cairns
;  APPLICANT:  Hugh Jonathan Herdon
;  TITLE OF INVENTION:  NOVEL COMPOUNDS
;  FILE REFERENCE:  GP-30176-D1
;  CURRENT APPLICATION NUMBER: US/09/795,232
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/182,728
;  PRIOR FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  9818890.7
;  PRIOR FILING DATE:  1998-08-28
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-795-232-2

  Query Match             27.0%;  Score 1031.5;  DB 10;  Length 797;
  Best Local Similarity   33.9%;  Pred. No. 8.3e-87;
  Matches  222;  Conservative  112;  Mismatches  186;  Indels  135;  Gaps   14;

Qy     36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
Db    173 SVATVATQEDERGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223

Qy     93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
Db    224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVL 282

Qy    153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC------------------------PVVR-- 185
Db    283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy    186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
Db    343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy    233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
Db    402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy    293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
Db    462 TNATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy    353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
Db    522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy    413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
Db    539 ---------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy    473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
Db    578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy    529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
Db    638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy    589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
Db    698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 12
US-09-843-598-10
; Sequence 10, Application US/09843598
; Patent No. US20020010944A1
; GENERAL INFORMATION:
;  APPLICANT: Horvitz, H. Robert
;  APPLICANT:  Ranganathan, Rajesh
;  TITLE OF INVENTION: CeSERT GENES, PROTEINS, AND MODULATORY
;  TITLE OF INVENTION:  COMPOUNDS
;  FILE REFERENCE: 01997/525002
;  CURRENT APPLICATION NUMBER: US/09/843,598
;  CURRENT FILING DATE:  2001-04-26
;  PRIOR APPLICATION NUMBER: US 60/200,549
;  PRIOR FILING DATE: 2000-04-26
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 630
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-843-598-10

  Query Match             24.6%;  Score 937.5;  DB 10;  Length 630;
  Best Local Similarity   30.8%;  Pred. No. 3.1e-78;
  Matches  199;  Conservative  128;  Mismatches  207;  Indels  113;  Gaps   13;

Qy     54 ELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPL 113
Db     73 ELHQGERETWGKKVDFLLSVIGYAVDLGNVWRFPYICYQNGGGAFLLPYTIMAIFGGIPL 132

Qy    114 FFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQ 173
Db    133 FYMELALGQYHRNGCISIWRKICPIFKGIGYAICIIAFYIASYYNTIMAWALYYLISSFT 192

Qy    174 YPLPWSECPVVRN-GSVAVVEAE------CEKSSATTYFWYREALDI--SDSISESGGLN 224
Db    193 DQLPWTSCKNSWNTGNCTNYFSEDNITWTLHSTSPAEEFYTRHVLQIHRSKGLQDLGGIS 252

Qy    225 WKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHM 284
Db    253 WQLALCIMLIFTVIYFSIWKGVKTSGKVVWVTATFPYIILSVLLVRGATLPGAWRGVLFY 312

Qy    285 FTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTS 344
Db    313 LKPNWQKLLETGVWIDAAAQIFFSLGPGFGVLLAFASYNKFNNNCYQDALVTSVVNCMTS 372
```

```
Qy    345 VLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYME 404
Db    373 FVSGFVIFTVLGYMAEMRNED---------------------------------------- 393

Qy    405 MDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFL 464
Db    394 ----VSEVAKD-------------------AGPSLLFITYAEAIANMPASTFFAIIFFL 429

Qy    465 MLINLGLGSMIGTMAGITTPIIDTSKVP------KEMFTVGCCVFTFLVGLLFVQRSGNY 518
Db    430 MLITLGLDSTFAGLEGVITAVLD--EFPHVWAKRRERFVLAVVITCFFGSLVTLTFGGAY 487

Qy    519 FVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPL 578
Db    488 VVKLLEEYATGPAVLTVALIEAVAVSWFYGITQFCRDVKEMLGFSPGWFWRICWVAISPL 547

Qy    579 CMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVL 638
Db    548 FLLFI----ICSFLMSPPQLRLF-------QYNY-PYWSIILGYCIGTSSFICIPTYIAY 595

Qy    639 RHFHLLSDGSNTLSVSYKKARMMKDISNLEENDETRFILSKVPSEAP 685
Db    596 R--LIITPGT-------FKERIIKSIT------------PETPTEIP 621
```

RESULT 13
US-09-815-923-4
; Sequence 4, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
; APPLICANT: Gill, Sarjeet S.
; APPLICANT: Ross, Linda S.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1el
; TITLE OF INVENTION: Target Sites for Insecticides
; FILE REFERENCE: 02307O-093800US
; CURRENT APPLICATION NUMBER: US/09/815,923
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Manduca sexta
; FEATURE:
; OTHER INFORMATION: serotonin transporter
US-09-815-923-4

Query Match 24.3%; Score 926.5; DB 9; Length 587;
Best Local Similarity 31.5%; Pred. No. 2.9e-77;
Matches 203; Conservative 131; Mismatches 208; Indels 103; Gaps 14;

```
Qy     47 KQKAVEEEL-DAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVL 105
Db     24 KSRSVVVSLTPARQRETWAKKAEFLLAVVGFAVDLGNVWRFPYICYQNGGGAFLIPYCVM 83

Qy    106 LIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSI 165
Db     84 LLFGGLPLFFLELALGQYHRCGCLTLWKRICPALKGVGYAICMIDIYMGMYYNTIIGWAV 143

Qy    166 FYF---FKSFQYPLPWSECPVVRNGSVA--VVEAECEKSSAT--TYFWYREALD--ISDS 216
Db    144 YYLIASLASINSVLPWTSCDNEWNTPLCTPVTSPQTNPNSSTPAKEFFERNVLEQHKSNG 203

Qy    217 ISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRG 276
Db    204 LDDMGPIKPSLALCVFGVFVLVYFSLWKGVRSAGKVVWVTALAPYVVLLILLARGVTLPG 263

Qy    277 AVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALV 336
Db    264 ATEGIRYYLTPEWHKLQNSKVWIDAASQIFFSLGPGFGTLLALSSYNKFNNNCYRDALIT 323

Qy    337 SFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSH 396
Db    324 SSINCLTSFLAGFVIFSVLGYMAHVQNK------------------------------- 351

Qy    397 LTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSP 456
Db    352 --------------SIEE-----VGL-------------EGPGLVFIVYPEAIATMTGSV 379

Qy    457 FWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKVP------KEMFTVGCCVFTFLVGLL 510
Db    380 FWAIIFFLMLITLGLDSTFGGLEAVTTALCD--EYPRVLGRHREVFVAVLLLFIYICALP 437

Qy    511 FVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFY 570
Db    438 TTTYGGVYLVDLLNVYGPGLAILFVVFAEAAGVCWVYGVDRFSEDVRTMLGHTPGWFWRT 497

Qy    571 MWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMAL--LITLIVVA 628
Db    498 CWSYISPVFLLVLFVFSVL------------AHEEMLGGEYTYPSWSITVGWVMTGTTVS 545

Qy    629 TLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDISNLEENDET 673
Db    546 CIPLYIIYKL----LITPGN---CINRIKTIQRPEVTSIPPADST 583
```

RESULT 14
US-09-843-598-11
; Sequence 11, Application US/09843598
; Patent No. US20020010944A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Ranganathan, Rajesh
; TITLE OF INVENTION: CeSERT GENES, PROTEINS, AND MODULATORY
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 01997/525002
; CURRENT APPLICATION NUMBER: US/09/843,598
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/200,549
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-843-598-11

Query Match 24.2%; Score 924; DB 10; Length 622;
Best Local Similarity 31.2%; Pred. No. 5.4e-77;
Matches 205; Conservative 117; Mismatches 228; Indels 108; Gaps 11;

```
Qy      4 NSKVTQREHSSEHVTES-VADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDRPA 62
Db     32 NNEDDSNEDDGDHTTPAKVTDPL--------------APKLANNERILVVSVTERTRET 76

Qy     63 WNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQ 122
Db     77 WGQKAEFLLAVIGFAVDLGNVWRFPYICYQNGGGAFLVPYCLFLIFGGLPLFYMELALGQ 136

Qy    123 RIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSEC- 181
Db    137 FHRCGCLSIWKRICPALKGVGYAICLIDIYMGMYYNTIIGWAVYYLFASFTSKLPWTSCD 196

Qy    182 -PVVRNGSVAVVEAECEK--SSATTYFWYREALDI--SDSISESGGLNWKMTLCLLVVWS 236
Db    197 NPWNTENCMQVTSENFTELATSPAKEFFERKVLESYKGNGLDFMGPVKPTLALCVFGVFV 256

Qy    237 IGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQ 296
Db    257 LVYFSLWKGVRSAGKVVWVTALAPYVVLIILLVRGVSLPGADEGIKYYLTPEWHKLKNSK 316

Qy    297 VWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLG 356
```

```
Db    317 VWIDAASQIFFSLGPGFGTLLALSSYNKFNNNCYRDALITSSINCLTSFLAGFVIFSVLG 376

Qy    357 FKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQ 416
          : | :
Db    377 YMAYVQ---------------------------------------------------------- 382

Qy    417 FSALGLDPCLLEDELDK-SVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMI 475
                    :  :||  ::| || || : ||:     | |||::|||||| ||| |
Db    383 -----------KTSIDKVGLEGPGLVFIVYPEAIATMSGSVFWSIIFFLMLITLGLDSTF 431

Qy    476 GTMAGITTPIIDTSKVP------KEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSAT 529
          | :  : | : |  : |      :|:| :    | ||  |  :   |   : |
Db    432 GGLEAMITALCD--EYPRVIGRRRELFVLLLLAFIFLCALPTMTYGGVVLVNFLNVYGPG 489

Qy    530 LPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASII 589
          | :  :| :|   | | ||  :|  :: :||| :|  |:   | ::||: :  :   ||:
Db    490 LAILFVVFVEAAGVFWFYGVDRFSSDVEQMLGSKPGLFWRICWTYISPVFLLTIFIFSIM 549

Qy    590 QLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDG 647
                  ||   | | |:|:|   :   :   : | ||:  : : |     |
Db    550 -----------GYKEMLGEEY-YYPDWSYQVGWAVTCSSVLCIPMYIIYKFFFASKGG 595


RESULT 15
US-09-815-923-14
; Sequence 14, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
;  APPLICANT: Gill, Sarjeet S.
;  APPLICANT:  Ross, Linda S.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1el
;  TITLE OF INVENTION:  Target Sites for Insecticides
;  FILE REFERENCE: 023070-093800US
;  CURRENT APPLICATION NUMBER: US/09/815,923
;  CURRENT FILING DATE:  2001-03-23
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 597
;   TYPE: PRT
;   ORGANISM: Manduca sexta
;   FEATURE:
;   OTHER INFORMATION: GABA transporter
US-09-815-923-14

  Query Match             23.9%;  Score 914;  DB 9;  Length 597;
  Best Local Similarity   29.7%;  Pred. No. 4.3e-76;
  Matches  205;  Conservative  122;  Mismatches  217;  Indels  146;  Gaps   15;

Qy     27 LEEPVDYKQSVLNVAGE-AGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWR 85
          :|   | :   : :: : :| |   |  : :  :| :| ||| :||: || ::||||:||
Db      1 METKNDSRSDDIELSAQGSGNKPSDVAVKSNLPERGSWASKLDFILSVIGLAIGLGNVWR 60

Qy     86 FPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFS 145
          ||||| ||||||:|:|| : | : |||:||:|||:|| :  | :||:  | |   |||::
Db     61 FPYLCYKNGGGAFLIPYFLTLFLAGIPMFFMELAMGQMLTIGGLGVFK-IAPIFKGIGYA 119

Qy    146 SCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVV---------------------- 184
          : ::  :: :|| ||: |:||||| | :  :|  |
Db    120 AAVMSCWMNVYYIVILAWAIFYFFMSMRSDVPWRNCDNYWNTATCVNPYDRKNLTCWSSL 179

Qy    185 -----------RNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLV 233
                     || | ||:      |     |  | || ||  |   | : |::  ||:
Db    180 GDMSTFCTLNGRNVSKAVL------SDPVKEFWERRALQISSGIEHIGNIRWELAGTLLL 233

Qy    234 VWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKML 293
          || :    : ||:: :|||:||::|||| :|   |:||: | ||::||     | : |:|
Db    234 VWVLCYFCIWKGVRWTGKVVYFTALFPYFLLTVLLIRGITLPGAMEGIKFYVMPNMSKLL 293

Qy    294 DPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFV 353
          : :|| :  ||:||  ||| | ::   |||| || : |  :|  :|  ||: |  |:|
Db    294 ESEVWIDAVTQIFFSYGLGLGTLVALGSYNKFTNNVYKDALIVCSVNSSTSMFAGFVIFS 353

Qy    354 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVK 413
          |:|| |:                                                |   : |
Db    354 VVGFMAH-----------------------------------------------EQQRPVAEV- 369

Qy    414 EDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGS 473
                              :  | ||||:|:  |:   | :| || :|| ||: :|| |
Db    370 --------------------AASGPGLAFLAYPSAVLQLPGAPLWSCLFFFMLLLIGLDS 409

Qy    474 MIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSAT 529
          || |  | :||      :  ||:|    |: ::|||| :   | |   : | |: :
Db    410 QFCTMEGFITAVIDEWPKLLRRRKEIFIAITCIISYLVGLSCISEGGMYVFQILDSYAVS 469

Qy    530 -LPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMW-KFVSPLCMAVLTTAS 587
           |  ::  | ::::| :|  :|   : ||:|: |  :: : |  |   :|::|    :
Db    470 GFCLLFLIFFECVSISWAFGVNRFYDGIKEMIGYYPTIWWKFCWVGFTPAICISVF-IFN 528

Qy    588 IIQLGVTPPAYSAWIKEEAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLR------ 639
          ::|  ||  |  :           :| |  |      |  :  :|  :::: |
Db    529 LVQW--TPIKYMNY----------EYPWWSHAFGWFTALSSMLCIPGYMIYLWRVTPGTW 576

Qy    640 --HFHLLSDGSNTLSVSYKKARMMKDISNL 667
            ||              |  |: :|: :|
Db    577 QEKFH------------KIVRIPEDVPSL 593


Search completed: April 21, 2003, 14:03:04
Job time : 25 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 13:54:43 ; Search time 29 Seconds
(without alignments)
737.602 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 3817
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3803	99.6	727	1	US-08-424-424B-2	Sequence 2, Appli
2	3803	99.6	727	5	PCT-US94-05363A-2	Sequence 2, Appli
3	1302.5	34.1	616	1	US-08-149-100-2	Sequence 2, Appli
4	1036.5	27.2	797	2	US-08-700-013B-19	Sequence 19, Appl
5	1036.5	27.2	797	4	US-09-191-468-124	Sequence 124, App
6	1031.5	27.0	797	2	US-08-700-013B-21	Sequence 21, Appl
7	1031.5	27.0	797	4	US-09-182-728A-2	Sequence 2, Appli
8	1031.5	27.0	797	4	US-09-795-232-2	Sequence 2, Appli
9	1029.5	27.0	797	4	US-09-191-468-122	Sequence 122, App
10	1028	26.9	799	2	US-08-700-013B-27	Sequence 27, Appl
11	1025.5	26.9	797	4	US-09-191-468-120	Sequence 120, App
12	1003	26.3	638	1	US-08-295-814E-13	Sequence 13, Appl
13	1003	26.3	638	1	US-08-240-783B-2	Sequence 2, Appli
14	1003	26.3	638	3	US-09-084-813-2	Sequence 2, Appli
15	1003	26.3	638	4	US-09-343-361-13	Sequence 13, Appl
16	1003	26.3	638	5	PCT-US92-09662-2	Sequence 2, Appli
17	995.5	26.1	687	3	US-08-834-467-2	Sequence 2, Appli
18	995.5	26.1	687	4	US-09-396-177-2	Sequence 2, Appli
19	985.5	25.8	617	1	US-07-879-617A-11	Sequence 11, Appl
20	985.5	25.8	617	1	US-08-301-722A-3	Sequence 3, Appli
21	985.5	25.8	617	1	US-08-240-783B-3	Sequence 3, Appli
22	985.5	25.8	617	1	US-08-753-985-11	Sequence 11, Appl
23	985.5	25.8	617	3	US-09-084-813-3	Sequence 3, Appli
24	985.5	25.8	617	5	PCT-US92-09662-3	Sequence 3, Appli
25	978.5	25.6	619	1	US-07-762-132A-2	Sequence 2, Appli
26	978.5	25.6	619	1	US-08-301-722A-4	Sequence 4, Appli
27	970.5	25.4	614	1	US-08-291-299-8	Sequence 8, Appli
28	970.5	25.4	614	5	PCT-US95-10579-8	Sequence 8, Appli
29	968.5	25.4	614	1	US-08-295-814E-12	Sequence 12, Appl
30	968.5	25.4	614	4	US-09-343-361-12	Sequence 12, Appl
31	966.5	25.3	599	1	US-08-295-814E-11	Sequence 11, Appl
32	966.5	25.3	599	1	US-08-240-783B-4	Sequence 4, Appli
33	966.5	25.3	599	3	US-09-084-813-4	Sequence 4, Appli
34	966.5	25.3	599	4	US-09-343-361-11	Sequence 11, Appl
35	966.5	25.3	599	5	PCT-US92-09662-4	Sequence 4, Appli
36	965.5	25.3	620	1	US-08-301-722A-2	Sequence 2, Appli
37	963.5	25.2	667	1	US-07-879-617A-8	Sequence 8, Appli
38	963.5	25.2	667	1	US-08-753-985-8	Sequence 8, Appli
39	955	25.0	602	1	US-08-291-299-9	Sequence 9, Appli
40	955	25.0	602	5	PCT-US95-10579-9	Sequence 9, Appli
41	955	25.0	614	1	US-08-543-881-2	Sequence 2, Appli
42	955	25.0	614	1	US-08-291-299-2	Sequence 2, Appli
43	955	25.0	614	5	PCT-US94-00119-2	Sequence 2, Appli
44	955	25.0	614	5	PCT-US95-10579-2	Sequence 2, Appli
45	951.5	24.9	599	1	US-08-301-722A-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-08-424-424B-2
; Sequence 2, Application US/08424424B
; Patent No. 5759854
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Neurotransmitter Transporter
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,424B
;      FILING DATE:  APRIL 21, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05363
;      FILING DATE:  MAY 25, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  727 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
US-08-424-424B-2

  Query Match             99.6%;  Score 3803;  DB 1;  Length 727;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches  725;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60

Qy     61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYKVPYLVLLIIIGIPLFFLELAV 120

Qy    121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy    181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CPVVRNESVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240

Qy    241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300

Qy    301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360

Qy    361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420

Qy    421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy    481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy    541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600

Qy    601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660

Qy    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720

Qy    721 STPESEL 727
          |||||||
Db    721 STPESEL 727


RESULT 2
PCT-US94-05363A-2
; Sequence 2, Application PC/TUS9405363A
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Neurotransmitter Transporter
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05363A
;      FILING DATE:  SUBMITTED HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-118
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  727 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
PCT-US94-05363A-2

  Query Match             99.6%;  Score 3803;  DB 5;  Length 727;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches  725;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60

Qy     61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYKVPYLVLLIIIGIPLFFLELAV 120

Qy    121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy    181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CPVVRNESVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240

Qy    241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300

Qy    301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360

Qy    361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420

Qy    421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy    481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy    541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600

Qy    601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660

Qy    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
```

```
Qy     721 STPESEL 727
           |||||||
Db     721 STPESEL 727


RESULT 3
US-08-149-100-2
; Sequence 2, Application US/08149100
; Patent No. 5559021
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kelli E.
;    APPLICANT:  Weinshank, Richard L.
;    TITLE OF INVENTION:  DNA ENCODING A NOVEL MAMMALIAN
;    TITLE OF INVENTION:  TRANSPORTER HOMOLOGOUS TO NEUROTRANSMITTER TRANSPORTERS
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/149,100
;      FILING DATE:  08-NOV-1993
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  44744/JPW/TEP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 977-9550
;      TELEFAX:  (212) 6640525
;      TELEX:  422523 COOP UI
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  616 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-149-100-2

  Query Match             34.1%;  Score 1302.5;  DB 1;  Length 616;
  Best Local Similarity   43.1%;  Pred. No. 1.3e-119;
  Matches  265;  Conservative  122;  Mismatches  191;  Indels   37;  Gaps   14;

Qy      38 LNVAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
           | :   |   |:   :|  | |    || | : ||:: | | ::|||||:||||||||
Db       3 LAIKRRASRGQRPGPDEKRARDMEKARPQWGNPLQFVFACISYAVGLGNVWRFPYLCQMY 62

Qy      94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
           |||::|||||::||: |:|| :|||||||:|:|||| |  | | | |:| :| :|  |:
Db      63 GGGSFLVPYLIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFL 122

Qy     154 GLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDI 213
            :|||||  |  :| | ||| ||||| ||:  | :    : ||||:|:| |||||: |:|
Db     123 SMYYNVINAWGFWYLFHSFQDPLPWSVCPL--NSNRTGYDEECEKASSTQYFWYRKTLNI 180

Qy     214 SDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLL 273
           | || |:||: |:  ||| : | :  : : :| :|:|||:||::| || ||  :|||||
Db     181 SPSIQENGGVQWEPALCLTLAWLMVYLCILRGTESTGKVVYFTALMPYCVLIIYLVRGLT 240

Qy     274 LRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDG 333
           | || :|:::|||||: :: :|: |   |||:|| |||||| :| |:|||:  |:|
Db     241 LHGATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNEPSNDCQKHA 300

Qy     334 ALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVN 393
           :|| ||  ||: |::| | : ||||    | |:     |::  | ||  | ||
Db     301 VIVSVINSSTSIFASIVTFSIYGFKATFNYENCL----NKVI-LLLTN--SFDL-----E 348

Qy     394 FSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTH 451
           ||  :  |: : : :   :::| :   :  | || ||: :|||||||||| : ||: :
Db     349 DGFLTASNLEEVKDYLASTYPNKYSEVFPHIRNCSLESELNTAVQGTGLAFIVYAEAIKN 408

Qy     452 FPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLV 507
           |  |||::| ||: ||:|||:|  | | ||: |    :| :|||  :   |:    |
Db     409 MEVSQLWSVLYFFMLLMLGMGSMLGNTAAILTPLTDSKVISSYLPKEAISGLVCLINCAV 468

Qy     508 GLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRF 567
           |::|   :|||:  :|:||:||| | |||::| ||| ::|| ::|  :|  | | ||  :
Db     469 GMVFTMEAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLRAMTG-RPLNW 527

Qy     568 YF-YMWKFVSPLC---MAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWP---MAL 620
           |:  || |||||    : :   :  | |     | ||   :| :  |   ::|   :|:
Db     528 YWKAMWAFVSPLLIIGLFIFYLSDYILTGTL--QYQAW---DATQGQLVTKDYPPHALAV 582

Qy     621 LITLIVVATLPIPVV 635
           :  |:  :|: ||:|
Db     583 IGLLVASSTMCIPLV 597


RESULT 4
US-08-700-013B-19
; Sequence 19, Application US/08700013B
; Patent No. 5919653
;  GENERAL INFORMATION:
;    APPLICANT:  Albert, Vivian R.
;    APPLICANT:  Kowalski, Leslie R.Z.
;    APPLICANT:  Borden, Laurence A.
;    APPLICANT:  McKelvy, Jeffrey F.
;    TITLE OF INVENTION:  Human Glycine Transporter
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,013B
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  317743-108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  797 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-700-013B-19

  Query Match             27.2%;  Score 1036.5;  DB 2;  Length 797;
  Best Local Similarity   34.0%;  Pred. No. 3.3e-93;
  Matches  223;  Conservative  111;  Mismatches  186;  Indels  135;  Gaps   14;
```

```
Qy      36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
           ||  ||    | | : ||         |  |:||| :||: :|:::|||||:|||||  :
Db     173 SVATVATQEDEQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223

Qy      93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
           |||||:|:|||::| : |:|:||||:::||   :| : ||  | | | | | :  |: :
Db     224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVL 282

Qy     153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC-----------------------PVVR-- 185
           : :|||||| :::|| | ||   |||  |                         | ::
Db     283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy     186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
                    | ::    ::  |   | :  || |   | ||  |   | : | : |||
Db     343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy     233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
           : | |   :: |||::||||:||:: ||||||   |:|||: | ||  || :  |||  |:
Db     402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy     293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
            |  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     462 TDATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy     353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
           |:|| ||    |  :||
Db     522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
                       : || |:||: : ||:|  | ||||:::|||||: |||
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy     473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
           :|  |:  | | | |      :  | :||:|||:  |::|  : : | |   : | |:|
Db     578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy     529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
           :  | :| | | : ::::|| ::| :::  |:||:|  |:   | ||:|  :  :   |
Db     638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy     589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
           |    |  | ::       ||   ||| | |   ::  : : ||::||::  ||
Db     698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739
```

RESULT 5
US-09-191-468-124
; Sequence 124, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 124
; LENGTH: 797
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-124

Query Match 27.2%; Score 1036.5; DB 4; Length 797;
Best Local Similarity 34.0%; Pred. No. 3.3e-93;
Matches 223; Conservative 111; Mismatches 186; Indels 135; Gaps 14;

```
Qy      36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
           ||  ||    | | : ||         |  |:||| :||: :|:::|||||:|||||  :
Db     173 SVATVATQEDEQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223
```

```
Qy      93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
           |||||:|:|||::| : |:|:||||:::||   :| : ||  | | | | | :  |: :
Db     224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVL 282

Qy     153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC-----------------------PVVR-- 185
           : :|||||| :::|| | ||   |||  |                         | ::
Db     283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy     186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
                    | ::    ::  |   | :  || |   | ||  |   | : | : |||
Db     343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy     233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
           : | |   :: |||::||||:||:: ||||||   |:|||: | ||  || :  |||  |:
Db     402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy     293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
            |  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     462 TDATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy     353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
           |:|| ||    |  :||
Db     522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
                       : || |:||: : ||:|  | ||||:::|||||: |||
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy     473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
           :|  |:  | | | |      :  | :||:|||:  |::|  : : | |   : | |:|
Db     578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy     529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
           :  | :| | | : ::::|| ::| :::  |:||:|  |:   | ||:|  :  :   |
Db     638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy     589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
           |    |  | ::       ||   ||| | |   ::  : : ||::||::  ||
Db     698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739
```

RESULT 6
US-08-700-013B-21
; Sequence 21, Application US/08700013B
; Patent No. 5919653
; GENERAL INFORMATION:
; APPLICANT: Albert, Vivian R.
; APPLICANT: Kowalski, Leslie R.Z.
; APPLICANT: Borden, Laurence A.
; APPLICANT: McKelvy, Jeffrey F.
; TITLE OF INVENTION: Human Glycine Transporter
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,013B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  317743-108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  797 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-700-013B-21

  Query Match             27.0%;  Score 1031.5;  DB 2;  Length 797;
  Best Local Similarity   34.2%;  Pred. No. 1e-92;
  Matches  224;  Conservative  109;  Mismatches  187;  Indels  135;  Gaps   14;

Qy      36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
           ||  ||    | | : ||         |  |:||| :||: :|::|||||:|||||| :
Db     173 SVATVATQEDEQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223

Qy      93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
           |||||:|:|||::| : |:|:||||:::||   :| : ||  | | | | | :  |  :
Db     224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLINSVL 282

Qy     153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC-----------------------PVVR-- 185
           : :|||||| :::|| | ||   |||  |                       | ::
Db     283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy     186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
                    | ::    ::  |   | :  || |   | ||  |   | : | : |||
Db     343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy     233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
           : | |   :: |||::||||:||:: ||||||   |:||: | ||  || :  ||| |:
Db     402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy     293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
           |  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     462 TDATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy     353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
           |:|| ||    |  :||
Db     522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
                           : || |:||: : ||:|  | ||||:::|||||: |||
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy     473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
           :|  |:  | | | |      :  | :||:|||  |::|   : : | |   : | |:|
Db     578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCVCFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy     529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
           :  | :| | | : ::::|| ::| :::  |:|||:|  |:   | ||:|  :  :   |
Db     638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy     589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
           |    |  | ::       ||   ||| | |   ::  : : ||::||::  ||
Db     698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 7
US-09-182-728A-2
; Sequence 2, Application US/09182728A
; Patent No. 6238883
; GENERAL INFORMATION:
;  APPLICANT:  BROWN, ANTHONY
;  APPLICANT:  CHAPMAN, CONRAD GERALD
;  APPLICANT:  GLOGER, ISRAEL SIMON
;  APPLICANT:  EVANS, JOANNE RACHEL
;  APPLICANT:  CAIRNS, WILLIAM
;  APPLICANT:  HERDON, HUGH
;  TITLE OF INVENTION:  NOVEL COMPOUNDS
;  FILE REFERENCE:  GP-30176
;  CURRENT APPLICATION NUMBER: US/09/182,728A
;  CURRENT FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  9818890.7
;  PRIOR FILING DATE:  1998-08-28
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-182-728A-2

  Query Match             27.0%;  Score 1031.5;  DB 4;  Length 797;
  Best Local Similarity   33.9%;  Pred. No. 1e-92;
  Matches  222;  Conservative  112;  Mismatches  186;  Indels  135;  Gaps   14;

Qy      36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
           ||  ||    | | : ||         |  |:||| :||: :|::|||||:|||||| :
Db     173 SVATVATQEDERGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223

Qy      93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
           |||||:|:|||::| : |:|:||||:::||   :| : ||  | | | | | :  |: :
Db     224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVL 282

Qy     153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC-----------------------PVVR-- 185
           : :|||||| :::|| | ||   |||  |                       | ::
Db     283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy     186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
                    | ::    ::  |   | :  || |   | ||  |   | : | : |||
Db     343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy     233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
           : | |   :: |||::||||:||:: ||||||   |:||: | ||  || :  ||| |:
Db     402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy     293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
           :  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     462 TNATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy     353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
           |:|| ||    |  :||
Db     522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
                           : || |:||: : ||:|  | ||||:::|||||: |||
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy     473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
           :|  |:  | | | |      :  | :||:|||  |::|   : : | |   : | |:|
Db     578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy     529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
           :  | :| | | : ::::|| ::| :::  |:|||:|  |:   | ||:|  :  :   |
Db     638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy     589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
           |    |  | ::       ||   ||| | |   ::  : : ||::||::  ||
Db     698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 8
US-09-795-232-2
; Sequence 2, Application US/09795232
; Patent No. 6426405
; GENERAL INFORMATION:
;  APPLICANT:  Anthony M. Brown
;  APPLICANT:  Conrad Gerald Chapman
;  APPLICANT:  Israel Simon Gloger
```

```
;  APPLICANT:  Joanne Rachel Evans
;  APPLICANT:  William Cairns
;  APPLICANT:  Hugh Jonathan Herdon
;  TITLE OF INVENTION:  NOVEL COMPOUNDS
;  FILE REFERENCE:  GP-30176-D1
;  CURRENT APPLICATION NUMBER: US/09/795,232
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/182,728
;  PRIOR FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  9818890.7
;  PRIOR FILING DATE:  1998-08-28
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-795-232-2

  Query Match             27.0%;  Score 1031.5;  DB 4;  Length 797;
  Best Local Similarity   33.9%;  Pred. No. 1e-92;
  Matches  222;  Conservative  112;  Mismatches  186;  Indels  135;  Gaps   14;

Qy      36 SVLNVA---GEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQK 92
           ||  ||    | | : ||          | |:||| :||: :|::|||||:||||||  :
Db     173 SVATVATQEDERGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQ 223

Qy      93 NGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLF 152
           |||||:|:|||::| : |:|:||||:::||   :| : ||  | | | | | :  |: :
Db     224 NGGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVL 282

Qy     153 VGLYYNVIIGWSIFYFFKSFQYPLPWSEC-------------------------PVVR-- 185
           : :|||||| :::|| | ||   |||  |                         | ::
Db     283 IAIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIK 342

Qy     186 ----------NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLL 232
                    | ::    ::  |   | :  || |   | ||  |   | : | :  |||
Db     343 NSTFCMTAYPNVTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLF 401

Qy     233 VVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKM 292
           : | |   :: |||::||||:||:: ||||||   |:||: | ||  || :  |||  |:
Db     402 LAWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKL 461

Qy     293 LDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVF 352
           :  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     462 TNATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIF 521

Qy     353 VVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTV 412
           |:|| ||    |  :||
Db     522 SVIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     413 KEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLG 472
                       : || |:|| : ||:|  | ||||:::|||||: |||
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLD 577

Qy     473 SMIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSA 528
           :|  |:  | | | |      :  | :||:|||:  |::|   : : | |  : | |:|
Db     578 TMFATIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAA 637

Qy     529 TLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASI 588
           :  | :| | | : :::|| ::| :::  |:|||:|  |:   | ||:|  : :  |
Db     638 SYALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSF 697

Qy     589 IQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
            |    | | ::       ||   ||| | |   ::  : : ||::||::  ||
Db     698 YQW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 9
US-09-191-468-122
; Sequence 122, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
;  APPLICANT: Gallagher, Michael J.
;  APPLICANT: Burgess, Loyd R.
;  APPLICANT: Brunden, Kurt R.
;  TITLE OF INVENTION: Human Glycine Transporter Type 2
;  FILE REFERENCE: 12311US01
;  CURRENT APPLICATION NUMBER: US/09/191,468A
;  CURRENT FILING DATE: 1998-11-12
;  NUMBER OF SEQ ID NOS: 124
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: Human
US-09-191-468-122

  Query Match             27.0%;  Score 1029.5;  DB 4;  Length 797;
  Best Local Similarity   33.8%;  Pred. No. 1.6e-92;
  Matches  221;  Conservative  110;  Mismatches  190;  Indels  133;  Gaps   12;

Qy      36 SVLNVAGEAGGKQKAVEEELDAED--RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
           ||  || :         |:|   ||  |  |:||| :||: :|::|||||:||||||  :|
Db     173 SVATVATQ--------EDEQGDEDKARGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQN 224

Qy      94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
           ||||:|:|||::| : |:|:||||:::||   :| : ||  | | | | | :  |: : :
Db     225 GGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVLI 283

Qy     154 GLYYNVIIGWSIFYFFKSFQYPLPWSEC------------------------------ 181
           :|||||| :::|| | ||   |||  |
Db     284 AIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIKN 343

Qy     182 --------PVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLV 233
                  | |   :  :  :   | :  || |   | ||  |   | : | : ||| :
Db     344 STFCMTAYPNVTMVNFTSLANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLFL 402

Qy     234 VWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKML 293
           | |   :: |||::||||:||:: ||||||   |:||: | ||  || :  |||  |:|
Db     403 AWVIVYASLAKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKLT 462

Qy     294 DPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFV 353
           |  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     463 DATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIFS 522

Qy     354 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVK 413
           |:|| ||    |  :||
Db     523 VIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     414 EDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGS 473
                             : || |:|| : ||:|  | ||||:::|||||: ||| :
Db     539 --------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLDT 578

Qy     474 MIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSAT 529
           |  ::  | | | |      :  | :||:|||:  |::|   : : | |   : | |:|:
Db     579 MFASIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAAS 638

Qy     530 LPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASII 589
             | :| | | : ::::|| ::| :::  |:|||:|  |:   | ||:|  : :  |
Db     639 YALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSFY 698

Qy     590 QLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
           |    | | ::       ||   ||| | |   ::  : : ||::||::  ||
Db     699 QW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 10
US-08-700-013B-27
; Sequence 27, Application US/08700013B
; Patent No. 5919653
;  GENERAL INFORMATION:
;    APPLICANT:  Albert, Vivian R.
;    APPLICANT:  Kowalski, Leslie R.Z.
;    APPLICANT:  Borden, Laurence A.
;    APPLICANT:  McKelvy, Jeffrey F.
```

```
;    TITLE OF INVENTION:  Human Glycine Transporter
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,013B
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  317743-108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  799 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-700-013B-27

  Query Match             26.9%;  Score 1028;  DB 2;  Length 799;
  Best Local Similarity   32.2%;  Pred. No. 2.3e-92;
  Matches  228;  Conservative  118;  Mismatches  208;  Indels  154;  Gaps   16;

Qy      43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
           | | : ||          |  |:||| :||: :|::|||||:||||||  :|||||:|:||
Db     185 EQGDENKA---------RGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQNGGGAFLIPY 235

Qy     103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
           |::| : |:|:||||:::||   :| : ||  | | | | | :  |: : : :||||||
Db     236 LMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVLIAIYYNVIIC 294

Qy     163 WSIFYFFKSFQYPLPWSEC-------------------------PVVR------------ 185
           :::|| | ||   |||  |                         | ::
Db     295 YTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVIGDHPKIQIKNSTFCMTAYP 354

Qy     186 NGSVAVVEAECEK---SSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAV 242
           | ::    ::  |   |. :  || |   | ||  |   | : | :  || : | |   ::
Db     355 NLTMVNFTSQANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLPFCLFLAWVIVYASL 413

Qy     243 GKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVA 302
            |||::||||:||:: ||||||   |:||: | ||  || :  |||  |: |  ||:: |
Db     414 AKGIKTSGKVVYFTATFPYVVLVILLIRGVTLPGAGAGIWYFITPKWEKLTDATVWKDAA 473

Qy     303 TQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIM 362
           ||:|| |   :||:|  ||||| |||: |  :|:  |  ||: |  |:| |:|| ||
Db     474 TQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIFSVIGFMAN-- 531

Qy     363 NEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGL 422
             |  :||
Db     532 ERKVNIENV---------------------------------------------------- 540

Qy     423 DPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGIT 482
                   : || |:||: : ||:|  | ||||:::|||||: ||| :|  |:  |
Db     541 -----------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLDTMFATIETIV 589

Qy     483 TPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVIL 538
           | | |       :  | :||:|||:  |::|   : : | |   : | |:|:  | :| |
Db     590 TSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAASYALVIIAIF 649

Qy     539 ENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAY 598
           | : ::::|| ::| :::  |:||:|  |:  | ||:|  :  :  | |   | |  |
Db     650 ELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSFYQW--EPMTY 707

Qy     599 SAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKA 658
            ::       ||   ||| | |   ::  : : ||::||:: :  |: |
Db     708 GSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIKMY--LAPG---------- 744

Qy     659 RMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETS 706
           | :          :  : : |  |   |  : ||        || ||
Db     745 RFI---------ERLKLVCSPQPDWGPF-LAQHRGERYKNMIDPLGTS 782


RESULT 11
US-09-191-468-120
; Sequence 120, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
;  APPLICANT: Gallagher, Michael J.
;  APPLICANT: Burgess, Loyd R.
;  APPLICANT: Brunden, Kurt R.
;  TITLE OF INVENTION: Human Glycine Transporter Type 2
;  FILE REFERENCE: 12311US01
;  CURRENT APPLICATION NUMBER: US/09/191,468A
;  CURRENT FILING DATE: 1998-11-12
;  NUMBER OF SEQ ID NOS: 124
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 120
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: Human
US-09-191-468-120

  Query Match             26.9%;  Score 1025.5;  DB 4;  Length 797;
  Best Local Similarity   33.6%;  Pred. No. 4.1e-92;
  Matches  220;  Conservative  110;  Mismatches  191;  Indels  133;  Gaps   12;

Qy      36 SVLNVAGEAGGKQKAVEEELDAED--RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
           ||  || :         |:|   ||  |  |:||| :||: :|::|||||:||||||  :|
Db     173 SVATVATQ--------EDEQGDEDKARGNWSSKLDFILSMVGYAVGLGNVWRFPYLAFQN 224

Qy      94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
           ||||:|:|||::| : |:|:|||||:::||   :| : ||  | | | | | :  |: : : :
Db     225 GGGAFLIPYLMMLALAGLPIFFLEVSLGQFASQGPVSVWKAI-PALQGCGIAMLIISVLI 283

Qy     154 GLYYNVIIGWSIFYFFKSFQYPLPWSEC------------------------------ 181
           :|||||| :::|| | ||   |||  |
Db     284 AIYYNVIICYTLFYLFASFVSVLPWGSCNNPWNTPECKDKTKLLLDSCVISDHPKIQIKN 343

Qy     182 --------PVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLV 233
                   | |    :   :   | :  || |   | ||  |   | : | : ||| :
Db     344 STFCMTAYPNVTMVNFTSLANKTFVSGSEEYFKY-FVLKISAGIEYPGEIRWPLALCLFL 402

Qy     234 VWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKML 293
           | |   :: |||::||||:||:: ||| ||   |:||: | ||  || :  |||  |:
Db     403 AWVIVYASLAKGIKTSGKVVYFTATFPYAVLVILLIRGVTLPGAGAGIWYFITPKWEKLT 462

Qy     294 DPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFV 353
           |  ||:: |||:|| |   :||:|  |||||  |||: |  :|:  |  ||: |  |:|
Db     463 DATVWKDAATQIFFSLSAAWGGLITLSSYNKFHNNCYRDTLIVTCTNSATSIFAGFVIFS 522

Qy     354 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVK 413
           |:|| ||    |  :||
Db     523 VIGFMAN--ERKVNIENV---------------------------------------- 538

Qy     414 EDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGS 473
                         : || |:||: : ||:|  | ||||:::|||||: ||| :
Db     539 -------------------ADQGPGIAFVVYPEALTRLPLSPFWAIIFFLMLLTLGLDT 578
```

```
Qy    474 MIGTMAGITTPIIDT----SKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSAT 529
Db    579 MFASIETIVTSISDEFPKYLRTHKPVFTLGCCICFFIMGFPMITQGGIYMFQLVDTYAAS 638

Qy    530 LPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASII 589
Db    639 YALVIIAIFELVGISYVYGLQRFCEDIEMMIGFQPNIFWKVCWAFVTPTILTFILCFSFY 698

Qy    590 QLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHL 643
Db    699 QW--EPMTYGSY-------RY---PNWSMVLGWLMLACSVIWIPIMFVIK-MHL 739


RESULT 12
US-08-295-814E-13
; Sequence 13, Application US/08295814E
; Patent No. 5658786
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kelli E.
;    APPLICANT:  Borden, Laurence A.
;    APPLICANT:  Hartig, Paul R.
;    APPLICANT:  Weinshank, Richard L.
;    TITLE OF INVENTION:  DNA ENCODING TAURINE AND GABA
;    TITLE OF INVENTION:  TRANSPORTERS AND USES THEREOF
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,814E
;      FILING DATE:  DECEMBER 19, 1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  40558-B-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0525
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  638 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-295-814E-13

  Query Match             26.3%;  Score 1003;  DB 1;  Length 638;
  Best Local Similarity   31.3%;  Pred. No. 4.8e-90;
  Matches  224;  Conservative  132;  Mismatches  224;  Indels  136;  Gaps   16;

Qy     40 VAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
Db      8 VATSSPEQNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGG 67

Qy     96 GAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGL 155
Db     68 GAFMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGI 126

Qy    156 YYNVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-V 191
Db    127 YYNVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHL 186

Qy    192 VEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGK 251
Db    187 FNYTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGK 246

Qy    252 VMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGL 311
Db    247 VVYFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGC 306

Qy    312 GFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENA 371
Db    307 AWGGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN----------- 355

Qy    372 EKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDEL 431
Db    356 ----------------------------------------------HLGVDVSRVAD-- 366

Qy    432 DKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKV 491
Db    367 ----HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EV 420

Qy    492 PKEMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
Db    421 GNEWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIM 480

Qy    545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
Db    481 YIYGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN----- 533

Qy    605 EAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMK 662
Db    534 -----HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSR 584

Qy    663 DIS-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
Db    585 DWGPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 636


RESULT 13
US-08-240-783B-2
; Sequence 2, Application US/08240783B
; Patent No. 5756348
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kelli
;    APPLICANT:  Borden, Laurence A.
;    APPLICANT:  Branchek, Theresa
;    APPLICANT:  Hartig, Paul R.
;    APPLICANT:  Weinshank, Richard L.
;    TITLE OF INVENTION:  DNA ENCODING A GLYCINE TRANSPORTER AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/240,783B
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  1795/39875-A-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212)391-0525
```

```
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  638 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-240-783B-2

  Query Match             26.3%;  Score 1003;  DB 1;  Length 638;
  Best Local Similarity   31.3%;  Pred. No. 4.8e-90;
  Matches  224;  Conservative  132;  Mismatches  224;  Indels  136;  Gaps   16;

Qy      40 VAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
           ||  :  :  ||  |   :|    |  | ::::::|  :|::|||||:|||||||:|||
Db       8 VATSSPEQNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGG 67

Qy      96 GAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGL 155
           ||:: || ::|:  |||||:||: ||   :| :|||  | |   |:|:   :|  ::|:
Db      68 GAFMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGI 126

Qy     156 YYNVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-V 191
           ||||:|  : :||| |  : |||: |                      |   :|::: :
Db     127 YYNVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHL 186

Qy     192 VEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGK 251
               :::| :  :|    | :|| | : | :   :  || | | :  : : :|::||||
Db     187 FNYTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGK 246

Qy     252 VMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGL 311
           |:||:: ||||||    |||: | ||  ||::  |||  |:|: :|| : |:|:|: ||
Db     247 VVYFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGC 306

Qy     312 GFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENA 371
            :||:|  :|||| |||: |  ::|  |  ||| |  |:| :||| ||
Db     307 AWGGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN----------- 355

Qy     372 EKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDEL 431
                                                    ||:|   : |
Db     356 ----------------------------------------------HLGVDVSRVAD-- 366

Qy     432 DKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKV 491
               | ||||:|: ||:|  | || ||::|| ||| ||||:    :  : | |:|  :|
Db     367 ----HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EV 420

Qy     492 PKEMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
             |         |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  :::
Db     421 GNEWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIM 480

Qy     545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
           :||| : : |::  |||| |  |:   |:|||| :  :   ::||    |  |:
Db     481 YIYGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN----- 533

Qy     605 EAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMK 662
               :  :| | :|:  |: |  |  :|:  :| |      :||   |       :  :
Db     534 -----HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSR 584

Qy     663 DIS-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
           |    | |:   |:     |: |||       | |       :|  :| |: |
Db     585 DWGPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 636


RESULT 14
US-09-084-813-2
; Sequence 2, Application US/09084813
; Patent No. 6127131
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kelli
;    APPLICANT:  Borden, Laurence A.
;    APPLICANT:  Branchek, Theresa
;    APPLICANT:  Hartig, Paul R.
;    APPLICANT:  Weinshank, Richard L.
;    TITLE OF INVENTION:  DNA ENCODING A GLYCINE TRANSPORTER AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/084,813
;      FILING DATE:  26-MAY-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  1795/39875-AZ-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212)391-0525
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  638 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-084-813-2

  Query Match             26.3%;  Score 1003;  DB 3;  Length 638;
  Best Local Similarity   31.3%;  Pred. No. 4.8e-90;
  Matches  224;  Conservative  132;  Mismatches  224;  Indels  136;  Gaps   16;

Qy      40 VAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
           ||  :  :  ||  |   :|    |  | ::::::|  :|::|||||:|||||||:|||
Db       8 VATSSPEQNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGG 67

Qy      96 GAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGL 155
           ||:: || ::|:  |||||:||: ||   :| :|||  | |   |:|:   :|  ::|:
Db      68 GAFMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGI 126

Qy     156 YYNVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-V 191
           ||||:|  : :||| |  : |||: |                      |   :|::: :
Db     127 YYNVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHL 186

Qy     192 VEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGK 251
               :::| :  :|    | :|| | : | :   :  || | | :  : : :|::||||
Db     187 FNYTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGK 246

Qy     252 VMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGL 311
           |:||:: ||||||    |||: | ||  ||::  |||  |:|: :|| : |:|:|: ||
Db     247 VVYFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGC 306

Qy     312 GFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENA 371
            :||:|  :|||| |||: |  ::|  |  ||| |  |:| :||| ||
Db     307 AWGGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN----------- 355

Qy     372 EKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDEL 431
                                                    ||:|   : |
Db     356 ----------------------------------------------HLGVDVSRVAD-- 366

Qy     432 DKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKV 491
               | ||||:|: ||:|  | || ||::|| ||| ||||:    :  : | |:|  :|
Db     367 ----HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EV 420

Qy     492 PKEMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
             |         |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  :::
Db     421 GNEWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIM 480

Qy     545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
```

```
          :||| : : |::  |||| |  |:   |:||||  :  :   ::||    |  |:
Db    481 YIYGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN----- 533

Qy    605 EAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMK 662
          :  :| | :|:  |: |  |  :|:  :|. |      :||   |       :  :
Db    534 -----HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSR 584

Qy    663 DIS-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
          |    | |:   |:     |:  |||       | |        :|   :| |: |
Db    585 DWGPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 636


RESULT 15
US-09-343-361-13
; Sequence 13, Application US/09343361
; Patent No. 6225115
; GENERAL INFORMATION:
;  APPLICANT: Smith, Kelli E. et al
;  TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
;  TITLE OF INVENTION: Thereof
;  FILE REFERENCE: 40558-D
;  CURRENT APPLICATION NUMBER: US/09/343,361
;  CURRENT FILING DATE: 1999-06-30
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
;   LENGTH: 638
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-09-343-361-13

  Query Match             26.3%;  Score 1003;  DB 4;  Length 638;
  Best Local Similarity   31.3%;  Pred. No. 4.8e-90;
  Matches  224;  Conservative  132;  Mismatches  224;  Indels  136;  Gaps   16;

Qy     40 VAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
          ||  :  :  ||  |   :|    |  | ::::::|  :|::|||||:|||||| :|||
Db      8 VATSSPEQNGAVPSEATKKDQNLTRGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGG 67

Qy     96 GAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGL 155
          ||:: || ::|:  ||||||:||: ||   :| :|||  | |   |:|:   :|  ::|:
Db     68 GAFMFPYFIMLVFCGIPLFFMELSFGQFASQGCLGVWR-ISPMFKGVGYGMMVVSTYIGI 126

Qy    156 YYNVIIGWSIFYFFKSFQYPLPWSEC----------------------PVVRNGSVA-V 191
          ||||:|  : :||| |  : |||: |                      |   :|::: :
Db    127 YYNVVICIAFYYFFSSMTHVLPWAYCNNPWNTPDCAGVLDASNLTNGSRPTALSGNLSHL 186

Qy    192 VEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGK 251
          :::| :  :|    | :|| | : | :   :  || | | :  : : :|::||||
Db    187 FNYTLQRTSPSEEYWRLYVLKLSDDIGDFGEVRLPLLGCLGVSWVVVFLCLIRGVKSSGK 246

Qy    252 VMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGL 311
          |:||:: |||||    |||: | ||  ||::  |||  |:|: :|| : |:|:|: ||
Db    247 VVYFTATFPYVVLTILFVRGVTLEGAFTGIMYYLTPKWDKILEAKVWGDAASQIFYSLGC 306

Qy    312 GFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENA 371
           :||:|  :|||| |||: |  ::|  |  ||| |  |:| :||| ||
Db    307 AWGGLITMASYNKFHNNCYRDSVIISITNCATSVYAGFVIFSILGFMAN----------- 355

Qy    372 EKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDEL 431
                                                  ||:|   : |
Db    356 ----------------------------------------------HLGVDVSRVAD-- 366

Qy    432 DKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTSKV 491
          | ||||:|: ||:|  | || ||::|| ||| ||||:    :  : | |:|  :|
Db    367 ----HGPGLAFVAYPEALTLLPISPLWSLLFFFMLILLGLGTQFCLLETLVTAIVD--EV 420

Qy    492 PKEMF-------TVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVA 544
            |         |:|  |  ||:|:    ::| |:: : |:|:|:  | :|  :  :::
Db    421 GNEWILQKKTYVTLGVAVAGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVSIM 480

Qy    545 WIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKE 604
          :||| : : |::  |||| |  |:   |:||||  :  :   ::||    |  |:
```

```
Db    481 YIYGHRNYFQDIQMMLGFPPPLFFQICWRFVSPTIIFFILIFTVIQ--YRPITYN----- 533

Qy    605 EAAERYLYFPNWPMAL--LITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMK 662
          :  :| | :|:  |: |  |  :|:  :| |       :||   |       :  :
Db    534 -----HYQYPGWAVAIGFLMALSSVICIPLYALFQL----CRTDGDTLLQRLKNATKPSR 584

Qy    663 DIS-NLEENDETRFILSKVPSEAPSP------MPTHRS-----YLGPGSTSPLETS 706
          |    | |:   |:     |:  |||       | |        :|   :| |: |
Db    585 DWGPALLEHRTGRY----APTTTPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDS 636


Search completed: April 21, 2003, 13:59:16
Job time : 34 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 13:46:37 ; Search time 40 Seconds
(without alignments)
2421.830 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 3817
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3810	99.8	727	17	AAR88390	Human neurotransmi
2	3681	96.4	727	23	AAE21800	Human HIPHUM 00000
3	2522.5	66.1	730	22	AAG67159	Amino acid sequenc
4	2522.5	66.1	730	22	AAY72908	Human NTT7 protein
5	2522.5	66.1	730	23	ABB77168	Human GABA transpo
6	2522.5	66.1	730	23	AAE14404	Human neurotransmi
7	2483	65.1	729	23	AAO21215	Protein sequence i
8	1495.5	39.2	675	22	AAE05100	Drosophila melanog
9	1495.5	39.2	744	22	ABB61363	Drosophila melanog
10	1323	34.7	616	22	AAG64193	Human nerve mass-t
11	1318	34.5	647	20	AAW73376	Human HPDDV78 prot
12	1307.5	34.3	592	23	AAE14408	Human neurotransmi
13	1307.5	34.3	634	23	AAE13282	Human transporters
14	1305	34.2	662	22	ABB59777	Drosophila melanog
15	1302.5	34.1	616	17	AAW07635	Rat brain derived
16	1273	33.4	392	22	AAG64743	Human sodium neuro
17	1037.5	27.2	797	19	AAW56380	Human GlyT-2 trans
18	1037.5	27.2	797	21	AAB23323	Human glycine tran
19	1036.5	27.2	797	19	AAW56372	Amino acid sequenc
20	1036.5	27.2	797	19	AAW56373	Human GlyT-2 trans
21	1036.5	27.2	797	19	AAW56374	Human GlyT-2 trans
22	1036.5	27.2	797	19	AAW56376	Human GlyT-2 trans
23	1036.5	27.2	797	19	AAW56385	Human GlyT-2 trans
24	1036.5	27.2	797	19	AAW56368	Amino acid sequenc
25	1036.5	27.2	797	21	AAB09898	Human glycine tran
26	1036.5	27.2	797	21	AAB23320	Human glycine tran
27	1036.5	27.2	797	21	AAB23321	Human glycine tran
28	1036.5	27.2	797	21	AAB23322	Human glycine tran
29	1035.5	27.1	797	19	AAW56382	Human GlyT-2 trans
30	1035.5	27.1	797	19	AAW56384	Human GlyT-2 trans
31	1032.5	27.1	797	21	AAB23324	Human glycine tran
32	1032.5	27.1	797	21	AAB23325	Human glycine tran
33	1032.5	27.1	797	21	AAB23326	Human glycine tran
34	1031.5	27.0	797	19	AAW56377	Human GlyT-2 trans
35	1031.5	27.0	797	19	AAW56378	Human GlyT-2 trans
36	1031.5	27.0	797	19	AAW56369	Amino acid sequenc
37	1031.5	27.0	797	21	AAY82115	Human SC polypepti
38	1030.5	27.0	797	19	AAW56375	Human GlyT-2 trans
39	1030.5	27.0	797	19	AAW56383	Human GlyT-2 trans
40	1029.5	27.0	797	19	AAW56386	Human GlyT-2 trans
41	1029.5	27.0	797	21	AAB09897	Human glycine tran
42	1028.5	26.9	797	19	AAW56381	Human GlyT-2 trans
43	1025.5	26.9	797	19	AAW56379	Human GlyT-2 trans
44	1025.5	26.9	797	21	AAB09896	Human glycine tran
45	1025.5	26.9	797	21	AAB23327	Human glycine tran

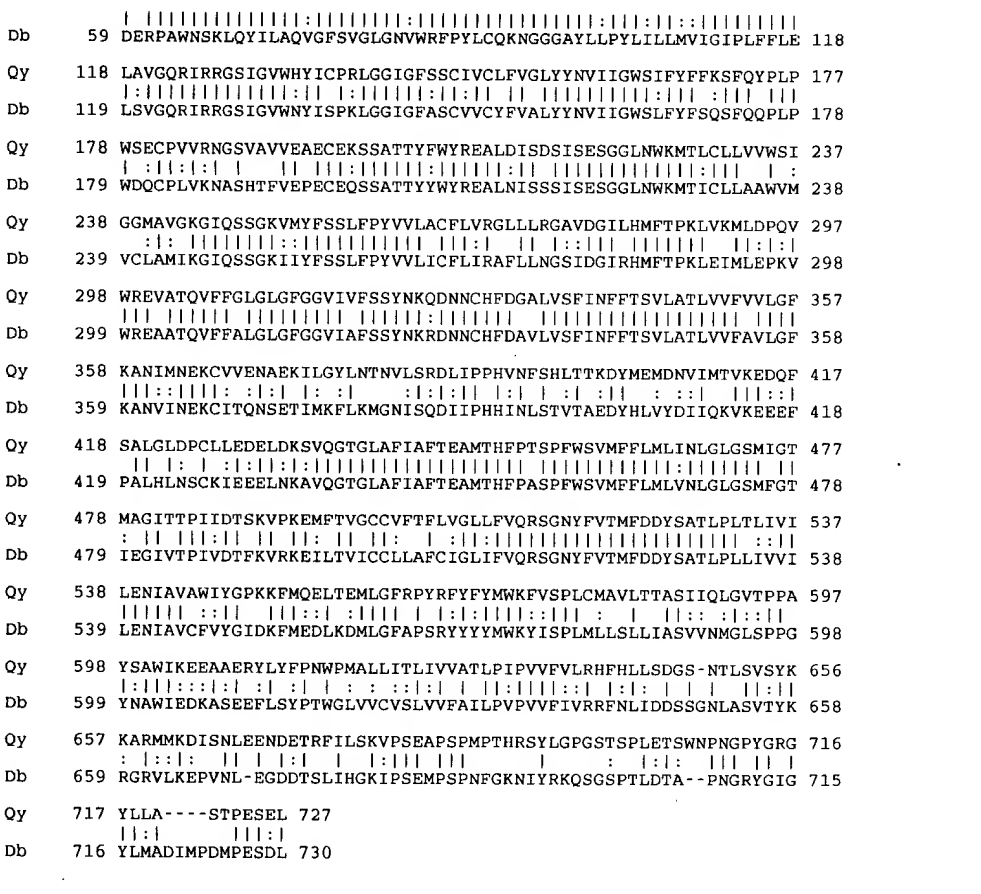
ALIGNMENTS

RESULT 1
AAR88390
ID AAR88390 standard; Protein; 727 AA.
XX
AC AAR88390;
XX
DT 15-OCT-1996 (first entry)
XX
DE Human neurotransmitter transporter protein.
XX
KW Neurotransmitter transporter protein; pain therapy; stroke therapy;
KW amyotrophic lateral sclerosis.
XX
OS Homo sapiens.
XX
PN WO9531539-A1.
XX
PD 23-NOV-1995.
XX
PF 16-MAY-1994; 94WO-US05363.
XX
PR 16-MAY-1994; 94WO-US05363.
PR 26-MAY-1994; 94ZA-0003696.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Fleischmann RD, Li Y;
XX
DR WPI; 1996-010925/01.
DR N-PSDB; AAT09866.
XX
PT DNA encoding neuro:transmitter transporter protein and related
PT (ant)agonists - useful to treat e.g., amyotrophic lateral sclerosis,

```
PT   pain and stroke.
XX
PS   Claim 1; Fig.1; 53pp; English.
XX
CC   This neurotransmitter transporter protein may be expressed
CC   recombinantly in a host microorganism using a vector plasmid.
CC   NTT antagonists may be used to treat amyotrophic lateral sclerosis,
CC   pain and strokes.  NTT antagonists/inhibitors may be used to treat
CC   neurological and psychiatric disorders, e.g. depression, anxiety
CC   and epilepsy.
XX
SQ   Sequence   727 AA;

  Query Match             99.8%;  Score 3810;  DB 17;  Length 727;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  726;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600

Qy     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 WIKEEAAEPYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660

Qy     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720

Qy     721 STPESEL 727
           |||||||.
Db     721 STPESEL 727


RESULT 2
AAE21800
ID   AAE21800 standard; Protein; 727 AA.
XX
AC   AAE21800;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Human HIPHUM 0000029 protein.
XX
KW   Human; neurotransmitter transporter polypeptide; HIPHUM 0000029 protein;
KW   prophylaxis; psychiatric disorder; bipolar disorder; psychotic disorder;
KW   unipolar depression; anxiety; schizophrenia; neurodegenerative disorder;
KW   neurological disorder; drug dependence; gene therapy; chromosome 1p21.
XX
OS   Homo sapiens.
XX
PN   GB2365432-A.
XX
PD   20-FEB-2002.
XX
PF   19-MAY-2000; 2000GB-0012199.
XX
PR   19-MAY-2000; 2000GB-0012199.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Sala CF,  Terstappen GC;
XX
DR   WPI; 2002-317841/36.
DR   N-PSDB; AAD34007.
XX
PT   An isolated neurotransmitter transporter polypeptide HIPHUM 0000029,
PT   whose regulation is useful in treating or preventing disease such as
PT   anxiety or depression -
XX
PS   Claim 1; Page 25-28; 39pp; English.
XX
CC   The invention relates to human neurotransmitter transporter polypeptides
CC   referred to as HIPHUM 0000029 and nucleic acid molecules encoding such
CC   polypeptides. Sequences of the invention their agonists and antagonists
CC   are useful for preparing a medicament for treatment or prophylaxis of
CC   disorders such as psychiatric disorders, bipolar disorders, unipolar
CC   depression, anxiety, schizophrenia, psychotic disorders, neurological/
CC   neurodegenerative disorders and drug dependence. Polynucleotides of the
CC   invention are useful as primers and probes. They are also used in gene
CC   therapy. The present sequence is human HIPHUM 0000029 protein. The
CC   HIPHUM 0000029 gene is located on chromosome 1p21.
XX
SQ   Sequence   727 AA;

  Query Match             96.4%;  Score 3681;  DB 23;  Length 727;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches  705;  Conservative    2;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKALEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVAWSIVGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
           || ||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     241 AVVKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLDKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
```

```
          |||||| |||||||| |||||||||||||| |||||||||||||||||||| |||||||
Db     301 AATQVFFALGLGFGGVIAFSSYNKQDNNCHFDAALVSFINFFTSVLATLVVFAVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMYNVIMTVKEDQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
           |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     481 ITTPIIDTFKVPKEMFTVGCCVFAFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
           |||||||| |:|||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     541 IAVAWIYGTKRFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPGYSA 600

Qy     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||| ||
Db     601 WIKEEAAERYLYFPNWAMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKGRM 660

Qy     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
           |||||||||||||||||||||||||||||||||||||||||||||||| |||| || |||||
Db     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSGNPNGRYGSGYLLA 720

Qy     721 STPESEL 727
           |||||||
Db     721 STPESEL 727
```

```
RESULT 3
AAG67159
ID   AAG67159 standard; Protein; 730 AA.
XX
AC   AAG67159;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human 579 transporter polypeptide.
XX
KW   Human; transporter; 20685; 579; 17114; 23821; 33894; 32613;
KW   vesicular monoamine transporter; neurotransmitter-symporter;
KW   ABC transporter; sulfate transporter; neurological disorder;
KW   central nervous system disorder; Parkinson's disease; depression; pain;
KW   infectious disease; cell proliferative disorder; cancer; blood disorder;
KW   immune disorder; inflammatory disorder; spleen disorder; lung disorder;
KW   Hodgkin's disease; Niemann-Pick disease; chronic bronchitis; ischemia;
KW   colon disorder; cirrhosis; uterus disorder; endometrium disorder;
KW   endometrial stromal tumour; brain disorder; T-cell disorder; anemia;
KW   Sjogren syndrome; skin disorder; lupus erythematosus; heart disorder;
KW   haematopoietic stem cell; Alzheimer's disease; myocardial infarction;
KW   blood vessel; Kawasaki syndrome; red cell disorder; thymus disorder;
KW   B-cell disorder; kidney disorder; glomerulonephritis; breast disorder;
KW   testis disorder; thyroid disorder; Graves disease; pancreatitis;
KW   skeletal muscle disorder; tumour; pancreas disorder;
KW   small intestine disorder; celiac sprue.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          61..659
FT                   /note= "sodium neurotransmitter-symporter domain"
FT   Domain          70..87
FT                   /note= "transmembrane domain"
FT   Domain          98..117
FT                   /note= "transmembrane domain"
FT   Domain          140..164
FT                   /note= "transmembrane domain"
FT   Domain          228..244
FT                   /note= "transmembrane domain"
FT   Domain          253..275
FT                   /note= "transmembrane domain"
FT   Domain          306..323
FT                   /note= "transmembrane domain"
FT   Domain          334..358
FT                   /note= "transmembrane domain"
FT   Domain          458..479
FT                   /note= "transmembrane domain"
FT   Domain          496..513
FT                   /note= "transmembrane domain"
FT   Domain          527..550
FT                   /note= "transmembrane domain"
FT   Domain          575..594
FT                   /note= "transmembrane domain"
FT   Domain          617..639
FT                   /note= "transmembrane domain"
XX
PN   WO200164875-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06374.
XX
PR   29-FEB-2000; 2000US-0185906.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA;
XX
DR   WPI; 2001-550178/61.
DR   N-PSDB; AAH75186.
XX
PT   Novel human transporter polypeptides useful for treating and diagnosing
PT   Parkinson's disease, Hodgkin disease, glomerulonephritis, myocardial
PT   infarction, Grave's disease, Alzheimer's disease, anemia, asthma and
PT   tumours -
XX
PS   Claim 9; Fig 9A-C; 259pp; English.
XX
CC   The present sequence represents a human transporter polypeptide. The
CC   specification describes 20685, 579, 17114, 23821, 33894 or 32613 human
CC   transporter polypeptides. The 20685 transporter is similar to vesicular
CC   monoamine transporters. The 579 transporter is similar to
CC   neurotransmitter-symporters. The 17114 transporter is similar to ABC
CC   transporters. The 32613 transporter is similar to sulfate transporters.
CC   The transporter polypeptides and polynucleotides are useful for treating
CC   and diagnosing neurological and central nervous system disorders (e.g.
CC   Parkinson's disease, depression, pain), infectious disease, cell
CC   proliferative disorders (e.g., cancer), blood disorders, and immune and
CC   inflammatory disorders. They are also useful for treating and
CC   diagnosing disorders involving the spleen (e.g., Hodgkin disease,
CC   Niemann-Pick disease), lung (e.g., chronic bronchitis), colon
CC   (cirrhosis), uterus and endometrium (e.g., endometrial stromal tumours),
CC   brain (e.g., ischemia), T-cells (e.g., Sjogren syndrome), skin (lupus
CC   erythematosus), haematopoietic stem cells (e.g, Alzheimer's disease),
CC   heart (e.g., myocardial infarction), blood vessels (e.g., Kawasaki
CC   syndrome), red cells (e.g., anemias), disorders involving thymus,
CC   B-cells, kidney (e.g., glomerulonephritis), disorders involving breast,
CC   testis, epididymis, prostate, thyroid (e.g., Graves disease), disorders
CC   involving skeletal muscle (e.g, tumour), pancreas (e.g., pancreatitis),
CC   small intestine (e.g., celiac sprue), disorders related to reduced
CC   platelet number and ovary.
XX
SQ   Sequence   730 AA;

  Query Match             66.1%;  Score 2522.5;  DB 22;  Length 730;
  Best Local Similarity   64.4%;  Pred. No. 6.3e-250;
  Matches  473;  Conservative  114;  Mismatches  135;  Indels   13;  Gaps    8;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
           ||||||| :||   : ||||| ||||: |:  |  :| | | | |:   |   |||  : |
Db       1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy      59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
```



```
Db     59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177

Db    119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy    178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237

Db    179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy    238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297

Db    239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy    298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357

Db    299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy    358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417

Db    359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy    418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477

Db    419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy    478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537

Db    479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy    538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597

Db    539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy    598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656

Db    599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy    657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716

Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727

Db    716 YLMADIMPDMPESDL 730
```

RESULT 4
AAY72908
ID AAY72908 standard; Protein; 730 AA.
XX
AC AAY72908;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human NTT7 protein.
XX
KW Human; NTT7 protein; therapy; psychiatric disorder; anxiety; depression;
KW schizophrenia; phobia; panic disorder; obsessive compulsive disorder;
KW Parkinson's disease; central nervous system disorder; cerebroprotective;
KW neurological disorder; stroke; pain; neuropathic pain; sleep disorder;
KW tranquilliser; analgesic; neuroleptic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200119854-A2.
XX
PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-GB03541.
XX
PR 15-SEP-1999; 99GB-0021833.
PR 17-MAR-2000; 2000GB-0006545.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Hill J, Duckworth DM, Farmer M, Pangalos M;
XX
DR WPI; 2001-244780/25.
DR N-PSDB; AAD03039.
XX
PT Novel NTT7 polypeptide useful for treating anxiety, depression,
PT schizophrenia, phobia, Parkinson's disease, stroke, pain, and
PT psychiatric, panic, neurological, central nervous system, obsessive
PT compulsive and sleep disorders
XX
PS Claim 3; Page 20; 27pp; English.
XX
CC The present sequence is human NTT7 protein encoded by a cDNA. NTT7 is
CC thought to be a member of the neurotransmitter family of polypeptides.
CC NTT7 sequences are useful for treating psychiatric disorders, anxiety,
CC depression, schizophrenia, phobias, panic disorder, obsessive compulsive
CC disorder, Parkinson's disease, central nervous system disorders, stroke,
CC neurological disorders, pain, neuropathic pain, sleep disorders, and
CC diseases in which neurotransmitters are implicated. NTT7 sequences are
CC useful for screening antagonists and agonists of NTT7 and as vaccines
CC for inducing immunological response in a mammal. NTT7 sequences and its
CC antibodies are useful to configure screening methods for detecting the
CC effect of added compounds on the production of mRNA and polypeptide in
CC cells. NTT7 is useful in conventional low capacity screening methods and
CC also in high-throughput screening (HTS) formats and is useful for
CC identifying membrane bound or soluble receptors. NTT7 antibody is useful
CC to isolate or identify clones expressing NTT7 or to purify NTT7 by
CC affinity chromatography. NTT7 is useful as diagnostic reagent for
CC detecting mutations in the associated gene, and for chromosome
CC localisation studies and tissue expression studies.
XX
SQ Sequence 730 AA;

Query Match 66.1%; Score 2522.5; DB 22; Length 730;
Best Local Similarity 64.4%; Pred. No. 6.3e-250;
Matches 473; Conservative 114; Mismatches 135; Indels 13; Gaps 8;

```
Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58

Db      1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy     59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117

Db     59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177

Db    119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy    178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237

Db    179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy    238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297

Db    239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy    298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357

Db    299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy    358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417

Db    359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy    418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477

Db    419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478
```

```
Qy    478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
          : || |||:|| || ||: || ||:  | :||:|||||||||||||||||||||| ::||
Db    479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy    538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
          |||||| ::||  |||::| :|||| | |:|:||||::||| :  |  ||:: :|::||
Db    539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy    598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
          |:|||:::|:| :| :| | : : ::|:| | ||:||||::| |:|: | | |  ||:||
Db    599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy    657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
          : |::|:  || | |:|  |  |:||| |||      |     :  |:|:  ||| || |
Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727
          ||:|      |||:|
Db    716 YLMADIMPDMPESDL 730
```

RESULT 5
ABB77168
ID ABB77168 standard; Protein; 730 AA.
XX
AC ABB77168;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human GABA transporter protein.
XX
KW Human; GABA; transporter protein; drug screening; therapeutic target.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Region	19..21
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Region	19..22
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Region	25..28
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Region	49..52
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Region	55..58
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Region	85..99
FT		/label= Sodium:neurotransmitter_symporter_family_
FT		signature_1
FT	Region	126..129
FT		/label= cAMP-_and_cGMP-dependant_protein_kinase_
FT		phosphorylation_site
FT	Region	137..139
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Modified-site	142..147
FT		/note= "Gly is N-myristoylated"
FT	Modified-site	187..190
FT		/note= "Asn is N-glycosylated"
FT	Region	191..194
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	213..216
FT		/note= "Asn is N-glycosylated"
FT	Region	217..220
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	246..251
FT		/note= "Gly is N-myristoylated"
FT	Region	250..252
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Modified-site	276..279
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	277..282
FT		/note= "Gly is N-myristoylated"
FT	Region	287..289
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Modified-site	311..316
FT		/note= "Gly is N-myristoylated"
FT	Modified-site	315..320
FT		/note= "Gly is N-myristoylated"
FT	Modified-site	383..386
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	394..397
FT		/note= "Asn is N-glycosylated"
FT	Region	399..402
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Region	425..427
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Modified-site	438..443
FT		/note= "Gly is N-myristoylated"
FT	Modified-site	473..478
FT		/note= "Gly is N-myristoylated"
FT	Modified-site	477..482
FT		/note= "Gly is N-myristoylated"
FT	Region	489..491
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Region	522..525
FT		/label= Casein_kinase_II_phosphorylation_site
FT	Region	606..614
FT		/label= Tyrosine_kinase_phosphorylation_site
FT	Modified-site	618..623
FT		/note= "Gly is N-myristoylated"
FT	Modified-site	650..655
FT		/note= "Gly is N-myristoylated"
FT	Region	656..658
FT		/label= Protein_kinase_C_phosphorylation_site
FT	Region	696..699
FT		/label= cAMP-_and_cGMP-dependant_protein_kinase_
FT		phosphorylation_site

XX
PN US2002031800-A1.
XX
PD 14-MAR-2002.
XX
PF 21-DEC-2000; 2000US-0741149.
XX
PR 05-SEP-2000; 2000US-229529P.
XX
PA (LIZZ/) LI Z.
PA (CHAT/) CHATURVEDI K.
PA (ZHUS/) ZHU S.
PA (WOOD/) WOODAGE T.
PA (GUEG/) GUEGLER K.
PA (WEBS/) WEBSTER M.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX
PI Li Z, Chaturvedi K, Zhu S, Woodage T, Guegler K, Webster M;
PI Di Francesco V, Beasley EM;
XX
DR WPI; 2002-361179/39.
DR N-PSDB; ABL55856, ABL55857.
XX
PT New peptides and nucleic acid sequences related to the GABA transporter
PT subfamily, useful in developing diagnostic compositions, as well as in
PT drug screening, particularly as models for the development of human
PT therapeutic targets -
XX
PS Claim 1; Fig 2; 69pp; English.
XX
CC The sequence represents a novel human transporter protein of the GABA
CC transporter subfamily. The invention relates to a novel isolated human
CC peptide of the GABA transporter subfamily. The peptide and nucleic acids
CC are useful in developing human therapeutics and diagnostic compositions.
CC These are also useful in drug screening, particularly as models for the
CC development of human therapeutic targets.
XX
SQ Sequence 730 AA;

```
  Query Match             66.1%;  Score 2522.5;  DB 23;  Length 730;
  Best Local Similarity   64.4%;  Pred. No. 6.3e-250;
  Matches  473;  Conservative  114;  Mismatches  135;  Indels   13;  Gaps    8;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
           ||||||| :||   : ||||| |||: |:  |  :| | | | |:   |   |||  : |
Db       1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy      59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
           | ||||||||||||||:||||||||:||||||||||||||||:|||:||::|||||||||
Db      59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy     118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
           |:|||||||||||||:|| |:||||||:||:|| || |||||||||:||| :||| |||
Db     119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy     178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237
           | :||:|:| |   || |||:||||||:||||||:|| ||||||||||||||:||| | :
Db     179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy     238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297
             :|: ||||||||::|||||||||||| |||:|  || |::||| |||||||  ||:|:|
Db     239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy     298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357
           ||| |||||| |||||||| ||||||:|||||||  ||||||||||||||||| ||||
Db     299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy     358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417
           |||::|||: :|:| |: :|    :|:|:|| |:| | :| :||  : ::|  |||::|
Db     359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy     418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
             || |: | :|:||:|:|||||||||||||||||| ||||||||||:|||||| ||
Db     419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy     478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
           : || |||:|| || ||: || ||:  | :||:|||||||||||||||||||| ::||
Db     479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy     538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
           |||||| ::||  |||::| :|||| | |:|:|||||::||| :  |  ||:: :|::||
Db     539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy     598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
           |:|||:::|:| :| :| | : : ::|:| | ||:|||||::| |:|: | | |  ||:||
Db     599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy     657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
           : |::|:  || | |:|  |  |:|||| |||      |     :  |:|:  ||| || |
Db     659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy     717 YLLA----STPESEL 727
           ||:|    . |||:|
Db     716 YLMADIMPDMPESDL 730


RESULT 6
AAE14404
ID   AAE14404 standard; Protein; 730 AA.
XX
AC   AAE14404;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human neurotransmitter transporter, NTT-2.
XX
KW   Human; neurotransmitter transporter; NTT-2; transport disorder;
KW   neurological disorder; psychiatric disorder; diabetes mellitus;
KW   amyotrophic lateral sclerosis; Parkinson's disease; prostate cancer;
KW   cardiac disorder; angina; Alzheimer's disease; amnesia; epilepsy;
KW   sickle cell anaemia; stroke; Huntington's disease; psychiatric disorder;
KW   acute stress disorder; anorexia nervosa; transgenic animal;
KW   gene therapy; sodium:neurotransmitter symporter family; SNF.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          100..117
FT                   /label= Transmembrane_domain
FT   Domain          335..359
FT                   /label= Transmembrane_domain
FT   Domain          460..483
FT                   /label= Transmembrane_domain
FT   Domain          494..512
FT                   /label= Transmembrane_domain
FT   Domain          572..589
FT                   /label= Transmembrane_domain
FT   Domain          619..640
FT                   /label= Transmembrane_domain
XX
PN   WO200190148-A2.
XX
PD   29-NOV-2001.
XX
PF   17-MAY-2001; 2001WO-US16283.
XX
PR   19-MAY-2000; 2000US-205518P.
PR   22-JUN-2000; 2000US-213956P.
PR   28-JUN-2000; 2000US-215105P.
PR   14-JUL-2000; 2000US-218947P.
PR   27-JUL-2000; 2000US-228448P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sanjanwala MS,  Walia NK,  Tribouley CM,  Yue H,  Gandhi AR,  Ding L;
PI   Yao MG,  Lal P,  Baughn MR,  Hafalia A,  Elliott VS,  Patterson C;
PI   Rankumar J;
XX
DR   WPI; 2002-097640/13.
DR   N-PSDB; AAD23974.
XX
PT   Novel human neurotransmitter transporter polypeptides and
PT   polynucleotides for diagnosing, preventing or treating transport,
PT   neurological and psychiatric disorders and for identifying modulators
PT   of therapeutic use
XX
PS   Claim 1; Page 111-113; 123pp; English.
XX
CC   The present sequence is human neurotransmitter transporter (NTT)-2
CC   (Incyte ID No: 6881669CD1). The NTT-2 polypeptide contains
CC   sodium:neurotransmitter symporter family (SNF) signature
CC   sequences. The NTT polypeptide and polynucleotide
CC   are useful for diagnosis, treatment and prevention of transport,
CC   neurological and psychiatric disorders. Transport disorders
CC   include akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia,
CC   cystic fibrosis, Becker's muscular dystrophy, diabetes mellitus,
CC   diabetes insipidus, myasthenia gravis, myocarditis, Parkinson's
CC   disease, prostate cancer; cardiac disorders associated with transport
CC   include angina, bradyarrhythmia, dermatomyositis, polymyositis;
CC   neurological disorders associated with transport include Alzheimer's
CC   disease, amnesia, bipolar disorder, dementia, depression, epilepsy,
CC   Tourette's disorder, schizophrenia, and other disorders associated with
CC   transport include neurofibromatosis, sickle cell anaemia, Wilson's
CC   disease, cataracts, infertility, hyperglycaemia, hypoglycaemia, Graves'
CC   disease, goitre, Cushing's disease, hypercholesterolaemia and
CC   cystinuria. Neurological disorders treatable include epilepsy, stroke,
CC   Huntington's disease, dementia, and other extrapyramidal disorder,
CC   motor neuron disorders, prion disease including kuru, metabolic disease
CC   of the nervous system, and other developmental disorders of the central
CC   nervous system, neuromuscular disorders, metabolic, endocrine and toxic
CC   myopathies, periodic paralysis, mental disorders including mood and
CC   anxiety. Psychiatric disorders include acute stress disorder, alcohol
CC   dependence, anorexia nervosa, anxiety, obsessive-compulsive disorder,
CC   panic disorder and sleep disorder. The polynucleotide is useful for
```

CC creating knockin humanised animals or transgenic animals to model
CC human disease and to detect and quantify gene expression in biopsied
CC tissues in which expression of NTT is correlated with disease. The
CC polynucleotide is also useful for generating hybridisation probes useful
CC in mapping the naturally occurring genomic sequence and oligonucleotide
CC primers derived from it are useful to detect single nucleotide
CC polymorphisms. NTT, its fragments and antibodies are useful as elements
CC on a microarray which is useful to monitor or measure protein-protein
CC interactions, drug-target interactions and gene expression profiles.
CC Sequences of the NTT polypeptide are used to analyse the proteome of a
CC tissue or cell type. The polypeptide of the invention is also useful for
CC screening its agonist, antagonist, modulator or a compound that binds
CC to it.
XX
SQ Sequence 730 AA;

Query Match 66.1%; Score 2522.5; DB 23; Length 730;
Best Local Similarity 64.4%; Pred. No. 6.3e-250;
Matches 473; Conservative 114; Mismatches 135; Indels 13; Gaps 8;

```
Qy      1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
          ||||||| :||   : ||||| |||: |:  |  :| | | | |:   |   |||  : |
Db      1 MPKNSKVVKRE-LDDDVTESVKDLLSNEDAADDAFKTSELIVDGQE-EKDTDVEEGSEVE 58

Qy     59 D-RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLE 117
          | |||||||||||||||:||||||||:|||||||||||||||||:|||:||::|||||||||
Db     59 DERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLMVIGIPLFFLE 118

Qy    118 LAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLP 177
          |:|||||||||||||:|| |:||||||:||:|| || ||||||||||:||| :||| |||
Db    119 LSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWSLFYFSQSFQQPLP 178

Qy    178 WSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSI 237
          | :||:|:| |   || |||:||||||:||||||:|| ||||||||||||:|||  | :
Db    179 WDQCPLVKNASHTFVEPECEQSSATTYYWYREALNISSSISESGGLNWKMTICLLAAWVM 238

Qy    238 GGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQV 297
            :|: ||||||||::||||||||||| |||:|  || |::||| |||||||  ||:|:|
Db    239 VCLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAFLLNGSIDGIRHMFTPKLEIMLEPKV 298

Qy    298 WREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGF 357
          ||| |||||| |||||||| ||||||:|||||||  ||||||||||||||||| ||||
Db    299 WREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVVFAVLGF 358

Qy    358 KANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQF 417
          |||::||||: :|:|  |: :|    :|:|:|| |:| | :| :||  : ::|  |||::|
Db    359 KANVINEKCITQNSETIMKFLKMGNISQDIIPHHINLSTVTAEDYHLVYDIIQKVKEEEF 418

Qy    418 SALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
            || |: | :|:||:|:|||||||||||||||||| ||||||||||:|||||| ||
Db    419 PALHLNSCKIEEELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLVNLGLGSMFGT 478

Qy    478 MAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVI 537
          : || |||:|| || ||: || ||:  | :||:||||||||||||||||||| ::||
Db    479 IEGIVTPIVDTFKVRKEILTVICCLLAFCIGLIFVQRSGNYFVTMFDDYSATLPLLIVVI 538

Qy    538 LENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPA 597
          |||||| ::|| |||::| :|||| | |:|:||||::||| :  |  ||:: :|::||
Db    539 LENIAVCFVYGIDKFMEDLKDMLGFAPSRYYYYMWKYISPLMLLSLLIASVVNMGLSPPG 598

Qy    598 YSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NTLSVSYK 656
          |:|||:::|:| :| :| | : : ::|:| | ||:||||::| |:|: | | |  ||:||
Db    599 YNAWIEDKASEEFLSYPTWGLVVCVSLVVFAILPVPVVFIVRRFNLIDDSSGNLASVTYK 658

Qy    657 KARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRG 716
          : |::|:  || | |:|  |  |:||| |||      |     :  |:|:  ||| || |
Db    659 RGRVLKEPVNL-EGDDTSLIHGKIPSEMPSPNFGKNIYRKQSGSPTLDTA--PNGRYGIG 715

Qy    717 YLLA----STPESEL 727
          ||:|    . |||:|
Db    716 YLMADIMPDMPESDL 730
```

RESULT 7
AAO21215
ID AAO21215 standard; Protein; 729 AA.
XX
AC AAO21215;
XX
DT 19-JUL-2002 (first entry)
XX
DE Protein sequence identified by SwissProt Accession No: Q08469.
XX
KW Antiparkinsonian; nootropic; neuroprotective; tranquiliser; vulnerary;
KW cerebroprotective; human sodium-dependent neurotransmitter transporter;
KW hybridising; central; peripheral nervous system disease; brain injury;
KW cerebrovascular disease; Parkinson's disease; corticobasal degeneration;
KW motor neuron disease; dementia; multiple sclerosis; post-stroke;
KW traumatic brain injury; stroke; post-traumatic brain injury;
KW small-vessel cerebrovascular disease; Alzheimer's disease; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200229048-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-EP11440.
XX
PR 05-OCT-2000; 2000US-237689P.
XX
PA (FARB) BAYER AG.
XX
PI Kohler RH;
XX
DR WPI; 2002-426112/45.
XX
PT New human sodium-dependent neurotransmitter transporter polypeptide,
PT the regulation of which is useful for treating central or peripheral
PT nervous system diseases e.g. brain injuries, Parkinson's disease, and
PT dementia -
XX
PS Disclosure; Fig 3; 98pp; English.
XX
CC The invention relates to a purified human sodium-dependent
CC neurotransmitter transporter polypeptide comprising a sequence of 259
CC amino acids, given in the specification. A nucleic acid encoding the
CC protein is useful for detecting a polynucleotide encoding the protein in
CC a biological sample by hybridising the protein-encoding polynucleotide to
CC a nucleic acid material of a biological sample to form a hybridisation
CC complex, and detecting the hybridisation complex. The protein of the
CC invention and its encoding gene are useful for screening for agents which
CC decrease the activity of sodium-dependent neurotransmitter transporter
CC polypeptide by contacting the test compound with the protein or its gene
CC and detecting binding of the test compound to the protein or its gene. A
CC vector comprising the polynucleotide of the invention or a reagent that
CC modulates the activity of the protein or its gene is useful for the
CC preparation of a medicament for modulating the activity of sodium-
CC dependent neurotransmitter transporter in a disease. The reagent is
CC useful for treating diseases such as central or peripheral nervous system
CC disease, where the central nervous system disorders are selected from
CC brain injuries, cerebrovascular diseases and their consequences,
CC Parkinson's disease, corticobasal degeneration, motor neuron disease,
CC dementia, including multiple sclerosis, traumatic brain injury, stroke,
CC post-stroke, post-traumatic brain injury, small-vessel cerebrovascular
CC disease, and Alzheimer's disease. The polynucleotide of the invention can
CC be used in gene therapy. This sequence represents the protein sequence
CC identified by SwissProt Accession No: Q08469 relating to the sodium-
CC dependent neurotransmitter transporter protein of the invention.
XX
SQ Sequence 729 AA;

Query Match 65.1%; Score 2483; DB 23; Length 729;
Best Local Similarity 63.8%; Pred. No. 7.3e-246;
Matches 472; Conservative 110; Mismatches 134; Indels 24; Gaps 8;

```
Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVD--YKQSVLNVAGEAGGKQKAVEEELDAE 58
           ||||||| :|:   : | ||| |||: |: |:   |:| | |  :        |:: |||
Db       1 MPKNSKVVKRD-LDDDVIESVKDLLSNEDSVEDVSKKSELIVDVQE-------EKDTDAE 52

Qy      59 D-------RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111
           |       |||||||||||||||:||||||||:|||||||||||||||||:|||:||::|||
Db      53 DGSEVDDERPAWNSKLQYILAQVGFSVGLGNVWRFPYLCQKNGGGAYLLPYLILLLVIGI 112

Qy     112 PLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKS 171
           ||||||||||||||||||||:|| |:|||||:||:|| || ||||||||::||| :|
Db     113 PLFFLELSVGQRIRRGSIGVWNYISPKLGGIGFASCVVCYFVALYYNVIIGWTLFYFSQS 172

Qy     172 FQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCL 231
           || |||| :||:|:| |   :| ||||||||||:|||||| || |||||||||:|||| ||
Db     173 FQQPLPWDQCPLVKNASHTYIEPECEKSSATTYYWYREALAISSSISESGGLDWKMTGCL 232

Qy     232 LVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVK 291
           |  | :  :|: ||||||||:|||||||||||| |||:| ||| |::||| |||||||
Db     233 LAAWVMVCLAMIKGIQSSGKIMYFSSLFPYVVLICFLIRSLLLNGSIDGIRHMFTPKLEM 292

Qy     292 MLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVV 351
           ||:|:|||| |||||| |||||||||| ||||||:|||||||  ||||||||||||||||
Db     293 MLEPKVWREAATQVFFALGLGFGGVIAFSSYNKRDNNCHFDAVLVSFINFFTSVLATLVV 352

Qy     352 FVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMT 411
           | |||||||||:|||: :|:| ||  | |  :| |:|| |:|  | :|  :||  : ::|
Db     353 FAVLGFKANIVNEKCISQNSEMILKLLKTGNVSWDVIPRHINLSAVTAEDYHVVYDIIQK 412

Qy     412 VKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGL 471
           |||::|: | |  | :|||||:|:|||||||||||||||||| ||||||||||||||||
Db     413 VKEEEFAVLHLKACQIEDELNKAVQGTGLAFIAFTEAMTHFPASPFWSVMFFLMLINLGL 472

Qy     472 GSMIGTMAGITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLP 531
           ||| ||: || ||::|| || ||: || ||:  | :||:|||||||||||||||||||||
Db     473 GSMFGTIEGIITPVVDTFKVRKEILTVICCLLAFCIGLMFVQRSGNYFVTMFDDYSATLP 532

Qy     532 LTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQL 591
           | ::||||||||:::||  ||:::||:|||| | ::|:||||::||| :  |  |||: :
Db     533 LLIVVILENIAVSFVYGIDKFLEDLTDMLGFAPSKYYYYMWKYISPLMLVTLLIASIVNM 592

Qy     592 GVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGS-NT 650
           |::|| |:|||||:|:| :| :| | | :  :|:|:| ||:||||||:|  :|: | | |
Db     593 GLSPPGYNAWIKEKASEEFLSYPMWGMVVCFSLMVLAILPVPVVFVIRRCNLIDDSSGNL 652

Qy     651 LSVSYKKARMMKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPN 710
           ||:||: |::|:  || : |:   | |:||| ||      |     :  |:|:  ||
Db     653 ASVTYKRGRVLKEPVNL-DGDDASLIHGKIPSEMSSPNFGKNIYRKQSGSPTLDTA--PN 709

Qy     711 GPYGRGYLLA---STPESEL 727
           | || |||:|     |||:|
Db     710 GRYGIGYLMADMPDMPESDL 729
```

RESULT 8
AAE05100
ID AAE05100 standard; Protein; 675 AA.
XX
AC AAE05100;
XX
DT 18-SEP-2001 (first entry)
XX
DE Drosophila melanogaster dmNTT4 protein.
XX
KW Fruit fly; invertebrate symporter cell surface receptor; dmNTT4;
KW sodium/neurotransmitter family; SNF; biopesticide; therapy.
XX
OS Drosophila melanogaster.
XX
PN WO200149848-A2.
XX
PD 12-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35551.
XX
PR 30-DEC-1999; 99US-0173929.
PR 15-MAR-2000; 2000US-0189399.
PR 23-MAR-2000; 2000US-0191686.
PR 23-MAR-2000; 2000US-0191687.
PR 23-MAR-2000; 2000US-0191688.
PR 23-MAR-2000; 2000US-0191695.
XX
PA (GENO-) GENOPTERA LLC.
XX
PI Kellerman KA, Keegan KP, Ebens AJ, Torpey J;
XX
DR WPI; 2001-441879/47.
DR N-PSDB; AAD09680.
XX
PT Novel invertebrate symporter cell surface receptor proteins and nucleic
PT acid encoding the protein useful as pesticide or drug target and to
PT identify compounds that have utility as therapeutics or pesticides -
XX
PS Claim 18; Page 59-61; 71pp; English.
XX
CC The invention relates to invertebrate symporter cell surface receptors
CC of the sodium/neurotransmitter family (SNF) and nucleic acid molecules
CC encoding such receptors. The SNF protein is useful for detecting a
CC candidate compound especially a putative pesticidal or pharmaceutical
CC agent that interacts with an invertebrate symporter cell surface
CC receptor protein or its fragment. Insect or worm genetically modified
CC to express protein of the invention is useful for studying invertebrate
CC symporter cell surface receptor protein activity, by detecting the
CC phenotype caused by the expression or mis-expression of the protein in
CC the animal. Nucleic acids encoding the invertebrate receptor protein or
CC their fragments are useful as biopesticides. SNF nucleic acids are
CC useful for generating mutant phenotypes in an animal model or living
CC cells that are used to study the regulation of genes encoding the
CC proteins which are useful as pesticide or drug targets. The genetically
CC modified organisms or cells are useful in screening assays to identify
CC pesticides or therapeutics and thus are useful in the identification of
CC new drug targets, therapeutic agents, diagnostics and prognostics
CC useful in treatment of disorders associated with ion channels. The
CC nucleic acid molecules are also useful as hybridisation probes.
CC The present sequence is Drosophila melanogaster (dm) NTT4 SNF
CC homologue which is referred as dmNTT4.
XX
SQ Sequence 675 AA;

```
  Query Match             39.2%;  Score 1495.5;  DB 22;  Length 675;
  Best Local Similarity   44.0%;  Pred. No. 2.8e-144;
  Matches  279;  Conservative  124;  Mismatches  198;  Indels   33;  Gaps    6;

Qy      59 DRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLEL 118
             :| ||: |:|: |: ||:|||||||||||||||:|||||:|:|:::|:||: ||||| :||
Db      54 ERAAWSGKMQFFLSIIGYSVGLGNIWRFPYLCQQNGGGAFLIPFMVMLILEGIPLFLIEL 113

Qy     119 AVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPW 178
             :|||:| |::|||: | | ||||| |||| ||| |||||||  |  || | ||:|||||
Db     114 GIGQRMRLGALGVWNTIHPWLGGIGISSCIVTLFVALYYNVIITWVFFYLFNSFRYPLPW 173

Qy     179 SECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIG 238
           | ||:  ||:   :| || ||| ||||||| || : |: | ||| | : |||:: |:|
Db     174 SSCPL--NGTGFELE-ECAKSSETTYFWYRTTLDAAPSMDEPGGLKWWIVLCLMLSWTIV 230

Qy     239 GMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVW 298
             | |||||||||:||:|||||:||  | :||: ||||  |::||:|||: |:|:| ||
Db     231 FFIVMKGIQSSGKVVYFTSLFPYIVLTIFFIRGITLRGAGAGLMHMYTPKVEKLLEPTVW 290

Qy     299 REVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFK 358
           : |||||:  || || :| | |||   |||  |  |||  |  |:: |::|:| :||||
Db     291 LDAATQVFYSFGLAFGSLIAFGSYNTPKNNCVRDVLLVSVCNAVTAIYASVVIFAILGFK 350

Qy     359 ANIMNEKCVVENAE-----KILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVK 413
           | :  ::||  | |     |:|| ||                    |   :  :  :
Db     351 ATVNVDRCVASNTEILVKNKLLG-LN--------------------DTQGYEQAMSLLN 388
```

```
Qy    414 EDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGS 473
Db    389 GTELLRLQLSECSLAHELDNAAEGTGLAFIVFTQAIVELPGAPFWAVLFFTMLLSLGLGS 448

Qy    474 MIGTMAGITTPIID---TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATL 530
Db    449 QIGILEGMLCTLFDIDIIKRVKKQHVTGVVCLFCFIVGFIFCTGAGEYWLKMFDSFAGTI 508

Qy    531 PLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQ 590
Db    509 GLVVVALMEMIAVIFIYGHERFTEDIFQMTGYRPGRYWQWTWRYIGPVIMVCILVSSVVF 568

Qy    591 LGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNT 650
Db    569 MVIRNPTYGAWNADLGMIEQKSYPNWVMGIALSMILAGVLPMPIVFLMRSFQCLKVDLDI 628

Qy    651 LSVSYKKARMMKDISNLEENDETRFILSKVPSEA 684
Db    629 HQGSIRRNETTASTKEMIDNDDDNMSPDMPPQDS 662
```

RESULT 9
ABB61363
ID ABB61363 standard; Protein; 744 AA.
XX
AC ABB61363;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10881.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL05466.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 10881; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 744 AA;

Query Match 39.2%; Score 1495.5; DB 22; Length 744;
Best Local Similarity 44.0%; Pred. No. 3.3e-144;
Matches 279; Conservative 124; Mismatches 198; Indels 33; Gaps 6;

```
Qy     59 DRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLEL 118
Db     54 ERAAWSGKMQFFLSIIGYSVGLGNIWRFPYLCQQNGGGAFLIPFMVMLILEGIPLFLIEL 113

Qy    119 AVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPW 178
Db    114 GIGQRMRLGALGVWNTIHPWLGGIGISSCIVTLFVALYYNVIITWVFFYLFNSFRYPLPW 173

Qy    179 SECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIG 238
Db    174 SSCPL--NGTGFELE-ECAKSSETTYFWYRTTLDAAPSMDEPGGLKWWIVLCLMLSWTIV 230

Qy    239 GMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVW 298
Db    231 FFIVMKGIQSSGKVVYFTSLFPYIVLTIFFIRGITLRGAGAGLMHMYTPKVEKLLEPTVW 290

Qy    299 REVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFK 358
Db    291 LDAATQVFYSFGLAFGSLIAFGSYNTPKNNCVRDVLLVSVCNAVTAIYASVVIFAILGFK 350

Qy    359 ANIMNEKCVVENAE-----KILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVK 413
Db    351 ATVNVDRCVASNTEILVKNKLLG-LN--------------------DTQGYEQAMSLLN 388

Qy    414 EDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGS 473
Db    389 GTELLRLQLSECSLAHELDNAAEGTGLAFIVFTQAIVELPGAPFWAVLFFTMLLSLGLGS 448

Qy    474 MIGTMAGITTPIID---TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATL 530
Db    449 QIGILEGMLCTLFDIDIIKRVKKQHVTGVVCLFCFIVGFIFCTGAGEYWLKMFDSFAGTI 508

Qy    531 PLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQ 590
Db    509 GLVVVALMEMIAVIFIYGHERFTEDIFQMTGYRPGRYWQWTWRYIGPVIMVCILVSSVVF 568

Qy    591 LGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNT 650
Db    569 MVIRNPTYGAWNADLGMIEQKSYPNWVMGIALSMILAGVLPMPIVFLMRSFQCLKVDLDI 628

Qy    651 LSVSYKKARMMKDISNLEENDETRFILSKVPSEA 684
Db    629 HQGSIRRNETTASTKEMIDNDDDNMSPDMPPQDS 662
```

RESULT 10
AAG64193
ID AAG64193 standard; Protein; 616 AA.
XX
AC AAG64193;
XX
DT 19-OCT-2001 (first entry)
XX
DE Human nerve mass-transferring protein.
XX
KW Human; nerve mass-transferring protein.
XX
OS Homo sapiens.
XX
PN CN1287170-A.
XX
PD 14-MAR-2001.
XX
PF 08-SEP-1999; 99CN-0118725.
XX
PR 08-SEP-1999; 99CN-0118725.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Zhao Y, Zhang H;

```
XX
DR   WPI; 2001-398933/43.
DR   N-PSDB; AAH73688.
XX
PT   New human nerve mass-transferring protein and its code sequence -
XX
PS   Claim 2; Page 17-19 (disclosure); 24pp; Chinese.
XX
CC   The invention provides a cDNA sequence encoding a new human nerve
CC   mass-transferring protein. The invention also relates to
CC   the application of the polynucleotide and the polypeptide, and the
CC   method of producing the polynucleotide and the polypeptide. The
CC   present sequence is the protein of the invention.
XX
SQ   Sequence   616 AA;

  Query Match           34.7%;  Score 1323;  DB 22;  Length 616;
  Best Local Similarity   41.7%;  Pred. No. 1.4e-126;
  Matches  263;  Conservative  125;  Mismatches  199;  Indels   44;  Gaps   11;

Qy      25 LALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIW 84
Db       3 LAIKKPASCDPRAGAERAEAGAMEKA---------RPLWANSLQFVFACISYAVGLGNVW 53

Qy      85 RFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGF 144
Db      54 RFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGV 113

Qy     145 SSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTY 204
Db     114 ASVVVSFFLSMYYNVINAWAFWYLFHSFQDPLPWSVCPL--NGNHTGYDEECEKASSTQY 171

Qy     205 FWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVL 264
Db     172 FWYRKTLNISPSLQENGGVQWEPALCLLLAWLVVYLCILRGTESTGKVVYFTASLPYCVL 231

Qy     265 ACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNK 324
Db     232 IIYLIRGLTLHGATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNE 291

Qy     325 QDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLS 384
Db     292 PSNNCQKHAIIVSLINSFTSIFASIVTFSIYGFKATFNYENCL-----KKVSLLLTNTFD 346

Qy     385 RDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAF 442
Db     347 -------LEDGFLTASNLEQVKGYLASAYPSKYSEMFPQIKNCSLESELDTAVQGTGLAF 399

Qy     443 IAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTV 498
Db     400 IVYTEAIKNMEVSQLWSVLYFFMLLMLGIGSMLGNTAAILTPLTDSKIISSHLPKEAISG 459

Qy     499 GCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTE 558
Db     460 LVCLVNCAIGMVFTMEAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLKA 519

Qy     559 MLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQLGVTPPAYSAWIKEEAAERYLY--- 612
Db     520 MTGRAVSWYWKVMWAGVSPLLIVSLFVFYLSDYILTGTL--KYQAW---DASQGQLVTKD 574

Qy     613 FPNWPMALLITLIVVATLPIPV----VFVLR 639
Db     575 YPAYALAVIGLLVASSTMCIPLAALGTFVQR 605


RESULT 11
AAW73376
ID   AAW73376 standard; Protein; 647 AA.
XX
AC   AAW73376;
XX
DT   12-FEB-1999  (first entry)
XX
DE   Human HPDDV78 protein sequence.
XX
KW   HPDDV78; human; neurotransmitter; mutation detection; gene therapy;
KW   brain hypoxia; seizure; stroke.
XX
OS   Homo sapiens.
XX
PN   EP881290-A2.
XX
PD   02-DEC-1998.
XX
PF   15-APR-1998;   98EP-0302907.
XX
PR   08-DEC-1997;   97EP-0309887.
PR   27-MAY-1997;   97GB-0010906.
XX
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Brown AM,  Evans JR;
XX
DR   WPI; 1999-001396/01.
DR   N-PSDB; AAV08571.
XX
PT   New neurotransmitter HPDDV78 polypeptides and polynucleotides -
PT   useful as diagnostic reagents and for prevention and treatment of
PT   brain hypoxia and stroke
XX
PS   Claim 10; Page 19-20; 24pp; English.
XX
CC   This sequence is the human neurotransmitter HPDDV78 polypeptide of
CC   the invention. HPDDV78 polypeptides and polynucleotides are useful for
CC   diagnosing susceptibility to diseases by detecting mutations in the
CC   HPDDV78 gene using probes containing the HPDDV78 nucleotide sequence, and
CC   can diagnose diseases associated with HPDDV78 imbalance by determining
CC   HPDDV78 polypeptide expression levels. HPDDV78 polypeptides can
CC   be used to screen for agonists and antagonists which bind HPDDV78
CC   by observing the binding, or stimulation or inhibition of HPDDV78
CC   activity. These can be used in treatment to activate (agonist)
CC   or inhibit (antagonist) HPDDV78 activity, in addition to direct
CC   administration of antisense sequences to prevent expression, or HPDDV78
CC   polypeptides to treat conditions associated with a lack of HPDDV78
CC   protein. Gene therapy may also be used to affect endogenous HPDDV78
CC   polypeptide production. HPDDV78 antibodies are useful for inducing an
CC   immune response to immunise and prevent diseases, and for isolating
CC   HPDDV78 clones or purifying the polypeptide by affinity
CC   chromatography. HPDDV78 polypeptides can be administered directly or as a
CC   vaccine to inoculate against disease. Diseases diagnosed, prevented or
CC   treated include: brain hypoxia, trauma, seizure and stroke. The HPDDV78
CC   polypeptide is also useful for mapping the gene to a chromosome, allowing
CC   gene inheritance to be studied through linkage analysis.
XX
SQ   Sequence   647 AA;

  Query Match           34.5%;  Score 1318;  DB 20;  Length 647;
  Best Local Similarity   42.4%;  Pred. No. 4.9e-126;
  Matches  260;  Conservative  122;  Mismatches  187;  Indels   44;  Gaps   11;

Qy      43 EAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPY 102
Db      52 EAGAMEKA---------RPLWANSLQFVFACISYAVGLGNVWRFPYLCQMYGGGSFLVPY 102

Qy     103 LVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIG 162
Db     103 IIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFLSMYYNVINA 162

Qy     163 WSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGG 222
Db     163 WAFWYLFHSFQDPLPWSVCPL--NGNHTGCDEECEKASSTQYFWYRKTLNISPSLQENGG 220

Qy     223 LNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGIL 282
Db     221 VQWEPALCLLLAWLVVYLCILRGTESTGKVVYFTASLPYCVLIIYLIRGLTLHGATNGLM 280
```

```
Qy    283 HMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFF 342
          :|||||: :: :|: |   |||:|| |||||| :| |:|||:  |||     :|| || |
Db    281 YMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNEPSNNCQKHAIIVSLINSF 340

Qy    343 TSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDY 402
          ||: |::| | : |||| | |:  | :  | ||          :   || :
Db    341 TSIFASIVTFSIYGFKATFNYENCL-----KKVSLLLTNTFD-------LEDGFLTASNL 388

Qy    403 MEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSV 460
           ::   : :    ::| :   :  | || ||| :|||||||| :|||: :   |  |||
Db    389 EQVKGYLASAYPSKYSEMFPQIKNCSLESELDTAVQGTGLAFIVYTEAIKNMEVSQLWSV 448

Qy    461 MFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRSG 516
          ::| ||: ||:|||:|  | | ||: |    :| :|||: :   |:    :|::|   :|
Db    449 LYFFMLLMLGIGSMLGNTAAILTPLTDSKIISSHLPKEAISGLVCLVNCAIGMVFTMEAG 508

Qy    517 NYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVS 576
          ||:  :|:||:||| | |||::| ||| ::|| ::|  :|  | |     ::  ||  ||
Db    509 NYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLKAMTGRAVSWYWKVMWAGVS 568

Qy    577 PLCMA---VLTTASIIQLGVTPPAYSAWIKEEAAERYLY---FPNWPMALLITLIVVATL 630
          || :    |   :  |  |     | ||   :|::  |    :| : :|::  |:  :|:
Db    569 PLLIVSLFVFYLSDYILTGTL--KYQAW---DASQGQLVTKDYPAYALAVIGLLVASSTM 623

Qy    631 PIPV----VFVLR 639
           ||:     || |
Db    624 CIPLAALGTFVQR 636


RESULT 12
AAE14408
ID   AAE14408 standard; Protein; 592 AA.
XX
AC   AAE14408;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human neurotransmitter transporter, NTT-6.
XX
KW   Human; neurotransmitter transporter; NTT-6; transport disorder;
KW   neurological disorder; psychiatric disorder; diabetes mellitus;
KW   amyotrophic lateral sclerosis; Parkinson's disease; prostate cancer;
KW   cardiac disorder; angina; Alzheimer's disease; amnesia; epilepsy;
KW   sickle cell anaemia; stroke; Huntington's disease; psychiatric disorder;
KW   acute stress disorder; anorexia nervosa; transgenic animal;
KW   gene therapy; sodium:neurotransmitter symporter family; SNF.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          83..103
FT                   /label= Transmembrane_domain
FT   Domain          174..199
FT                   /label= Transmembrane_domain
FT   Domain          277..297
FT                   /label= Transmembrane_domain
FT   Domain          392..415
FT                   /label= Transmembrane_domain
FT   Domain          508..524
FT                   /label= Transmembrane_domain
XX
PN   WO200190148-A2.
XX
PD   29-NOV-2001.
XX
PF   17-MAY-2001; 2001WO-US16283.
XX
PR   19-MAY-2000; 2000US-205518P.
PR   22-JUN-2000; 2000US-213956P.
PR   28-JUN-2000; 2000US-215105P.
PR   14-JUL-2000; 2000US-218947P.
PR   27-JUL-2000; 2000US-228448P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sanjanwala MS,  Walia NK,  Tribouley CM,  Yue H,  Gandhi AR,  Ding L;
PI   Yao MG,  Lal P,  Baughn MR,  Hafalia A,  Elliott VS,  Patterson C;
PI   Rankumar J;
XX
DR   WPI; 2002-097640/13.
DR   N-PSDB; AAD23978.
XX
PT   Novel human neurotransmitter transporter polypeptides and
PT   polynucleotides for diagnosing, preventing or treating transport,
PT   neurological and psychiatric disorders and for identifying modulators
PT   of therapeutic use -
XX
PS   Claim 1; Page 117-118; 123pp; English.
XX
CC   The present sequence is human neurotransmitter transporter (NTT)-6
CC   (Incyte ID No: 7475553CD1). The NTT-6 polypeptide contains
CC   sodium:neurotransmitter symporter family (SNF) signature
CC   sequences. The NTT polypeptide and polynucleotide
CC   are useful for diagnosis, treatment and prevention of transport,
CC   neurological and psychiatric disorders. Transport disorders
CC   include akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia,
CC   cystic fibrosis, Becker's muscular dystrophy, diabetes mellitus,
CC   diabetes insipidus, myasthenia gravis, myocarditis, Parkinson's
CC   disease, prostate cancer; cardiac disorders associated with transport
CC   include angina, bradyarrhythmia, dermatomyositis, polymyositis;
CC   neurological disorders associated with transport include Alzheimer's
CC   disease, amnesia, bipolar disorder, dementia, depression, epilepsy,
CC   Tourette's disorder, schizophrenia, and other disorders associated with
CC   transport include neurofibromatosis, sickle cell anaemia, Wilson's
CC   disease, cataracts, infertility, hyperglycaemia, hypoglycaemia, Graves'
CC   disease, goitre, Cushing's disease, hypercholesterolaemia and
CC   cystinuria. Neurological disorders treatable include epilepsy, stroke,
CC   Huntington's disease, dementia, and other extrapyramidal disorder,
CC   motor neuron disorders, prion disease including kuru, metabolic disease
CC   of the nervous system, and other developmental disorders of the central
CC   nervous system, neuromuscular disorders, metabolic, endocrine and toxic
CC   myopathies, periodic paralysis, mental disorders including mood and
CC   anxiety. Psychiatric disorders include acute stress disorder, alcohol
CC   dependence, anorexia nervosa, anxiety, obsessive-compulsive disorder,
CC   panic disorder and sleep disorder. The polynucleotide is useful for
CC   creating knockin humanised animals or transgenic animals to model
CC   human disease and to detect and quantify gene expression in biopsied
CC   tissues in which expression of NTT is correlated with disease. The
CC   polynucleotide is also useful for generating hybridisation probes useful
CC   in mapping the naturally occurring genomic sequence and oligonucleotide
CC   primers derived from it are useful to detect single nucleotide
CC   polymorphisms. NTT, its fragments and antibodies are useful as elements
CC   on a microarray which is useful to monitor or measure protein-protein
CC   interactions, drug-target interactions and gene expression profiles.
CC   Sequences of the NTT polypeptide are used to analyse the proteome of a
CC   tissue or cell type. The polypeptide of the invention is also useful for
CC   screening its agonist, antagonist, modulator or a compound that binds
CC   to it.
XX
SQ   Sequence   592 AA;

  Query Match             34.3%;  Score 1307.5;  DB 23;  Length 592;
  Best Local Similarity   42.8%;  Pred. No. 5.1e-125;
  Matches  255;  Conservative  121;  Mismatches  185;  Indels   35;  Gaps   10;

Qy     60 RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELA 119
          || | : ||:: | | ::|||||:|||||||  |||::||||:::||: |:|| :||||
Db      5 RPLWANSLQFVFACISYAVGLGNVWRFPYLCQMYGGGSFLVPYIIMLIVEGMPLLYLELA 64

Qy    120 VGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWS 179
          ||||:|:|||| |  | | | |:| :| :|  |: :|||||  |: :| | ||| |||||
Db     65 VGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFLSMYYNVINAWAFWYLFHSFQDPLPWS 124

Qy    180 ECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGG 239
             ||:  ||:    : ||||:|:| |||||: |:|| |: |:||: |:  ||||: | :
Db    125 VCPL--NGNHTGYDEECEKASSTQYFWYRKTLNISPSLQENGGVQWEPALCLLLAWLVVY 182
```

```
Qy    240 MAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWR 299
          : : :| :|:|||:||::  || ||  :|:||| | || :|:::|||||: :: :|: |
Db    183 LCILRGTESTGKVVYFTASLPYCVLIIYLIRGLTLHGATNGLMYMFTPKIEQLANPKAWI 242

Qy    300 EVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKA 359
          |||:|| |||||| :| |:|||:  |||     :|| || |||: |::| | : ||||
Db    243 NAATQIFFSLGLGFGSLIAFASYNEPSNNCQKHAIIVSLINSFTSIFASIVTFSIYGFKA 302

Qy    360 NIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSA 419
            | |:     | :  | ||          :    ||  :  ::   : :    ::|
Db    303 TFNYENCL-----KKVSLLLTNTFD-------LEDGFLTASNLEQVKGYLASAYPSKYSE 350

Qy    420 L--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGT 477
          :   :  | || ||| :|||||||||| :|||: :   |  |||::| ||: ||:|||:|
Db    351 MFPQIKNCSLESELDTAVQGTGLAFIVYTEAIKNMEVSQLWSVLYFFMLLMLGIGSMLGN 410

Qy    478 MAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLT 533
          | | ||: |    :| :|||  :   |:    :|::|   :|||:  :|:||:||| |
Db    411 TAAILTPLTDSKIISSHLPKEAISGLVCLVNCAIGMVFTMEAGNYWFDIFNDYAATLSLL 470

Qy    534 LIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMA---VLTTASIIQ 590
          |||::| ||| ::|| ::|  :|  | |      ::  ||  |||| :    |   :  |
Db    471 LIVLVETIAVCYVYGLRRFESDLKAMTGRAVSWYWKVMWAGVSPLLIVSLFVFYLSDYIL 530

Qy    591 LGVTPPAYSAWIKEEAAERYLY---FPNWPMALLITLIVVATLPIPV----VFVLR 639
            |    | ||   :|::  |    :| : :|::  |:  :|: ||:     || |
Db    531 TGTL--KYQAW---DASQGQLVTKDYPAYALAVIGLLVASSTMCIPLAALGTFVQR 581
```

RESULT 13
AAE13282
ID AAE13282 standard; Protein; 634 AA.
XX
AC AAE13282;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human transporters and ion channels (TRICH)-9.
XX
KW Human; transporter and ion channel; TRICH; akinesia; cystic fibrosis;
KW diabetes mellitus; Parkinson's disease; myasthenia gravis; dementia;
KW cardiac disorder; angina; hypertension; myocarditis; hyperglycaemia;
KW neurological disorder; Alzheimer's disease; cataract; infertility;
KW Wilson's disease; schizophrenia; Grave's disease; addison's disease;
KW Huntington's disease; multiple sclerosis; meningitis; hypotensive;
KW cardiant; nootropic; neuroprotective; neuroleptic; ophthalmological;
KW antithyroid; anticonvulsant; goitre; antiinflammatory.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 70..88
FT /label= Transmembrane_domain
FT Domain 123..141
FT /label= Transmembrane_domain
FT Domain 305..324
FT /label= Transmembrane_domain
FT Domain 417..440
FT /label= Transmembrane_domain
FT Domain 462..480
FT /label= Transmembrane_domain
FT Domain 533..550
FT /label= Transmembrane_domain
XX
PN WO200177174-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US11206.
XX
PR 06-APR-2000; 2000US-195595P.
PR 12-APR-2000; 2000US-196872P.
PR 20-APR-2000; 2000US-199020P.
PR 28-APR-2000; 2000US-200552P.
PR 05-MAY-2000; 2000US-202348P.
PR 11-MAY-2000; 2000US-203495P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Reddy R, Thornton M, Borowsky ML, Tang YT, Khan FA, Tribouley CM;
PI Gandhi AR, Yao MG, Sanjanwala MS, Baughn MR, Nguyen DB;
PI Policky JL, Yue H, Seilhamer JJ, Walia NK, Lal P, Kearney L;
PI Walsh RT, Lu DAM, Lu Y, Greene BD, Raumann BE, Patterson C;
XX
DR WPI; 2002-017448/02.
DR N-PSDB; AAD22001.
XX
PT Polypeptides of human transporters and ion channels, useful for
PT diagnosing, treating or preventing disorders of transport,
PT neurological, muscle, immunological and cell proliferative disorders -
XX
PS Claim 1; Page 128-130; 150pp; English.
XX
CC The invention relates to human transporters and ion channels (TRICH)
CC and the polynucleotides encoding them. The composition comprising TRICH
CC or agonist of TRICH is useful for treating a disease or condition
CC associated with decreased expression of functional TRICH or condition
CC associated with overexpression of TRICH respectively. The composition
CC comprising Ab is useful for diagnosing a condition of disease associated
CC with expression of TRICH in a subject, where the disorders include a
CC transport disorder such as akinesia, cystic fibrosis, diabetes mellitus,
CC Parkinson's disease, myasthenia gravis, cardiac disorders associated
CC with transport e.g. angina, hypertension, myocarditis, neurological
CC disorders associated with transport e.g. Alzheimer's disease, Wilson's
CC disease, schizophrenia, cataracts, infertility, hyperglycaemia, Grave's
CC disease, goitre, addison's disease, Huntington's disease, dementia,
CC multiple sclerosis, bacterial and viral meningitis. TRICH DNA is useful
CC for generating a transcript image of a tissue or cell type, which
CC represents the global pattern of gene expression by a particular tissue
CC or cell type and for analysing the proteome of a tissue or cell type.
CC TRICH DNA is used in gene therapy. The present amino acid sequence is
CC human TRICH9 protein.
XX
SQ Sequence 634 AA;

Query Match 34.3%; Score 1307.5; DB 23; Length 634;
Best Local Similarity 41.5%; Pred. No. 5.7e-125;
Matches 259; Conservative 115; Mismatches 213; Indels 37; Gaps 10;

```
Qy     49 KAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLII 108
          : :|:| :|  || |::| ||:|  :|| ||||||:|||||||| :||||:::|:|:||::
Db     22 ETIEQE-EASSRPKWDNKAQYMLTCLGFCVGLGNVWRFPYLCQSHGGGAFMIPFLILLVL 80

Qy    109 IGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYF 168
            |||| :|| |:|||:|||||:|||  | | | |:| :| :    ||||||| || | ::|
Db     81 EGIPLLYLEFAIGQRLRRGSLGVWSSIHPALKGLGLASMLTSFMVGLYYNTIISWIMWYL 140

Qy    169 FKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMT 228
          | ||| |||||: |:  | :  |   || :||   |||||| |:|| |||:|| : | |
Db    141 FNSFQEPLPWSDWPLNENQTGYV--DECARSSPVDYFWYRETLNISTSISDSGSIQWWML 198

Qy    229 LCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPK 288
          |||   ||:  |   :||:::|| :| :|  |||||  ||:||| |:|| :||: :|||
Db    199 LCLACAWSVLYMCTIRGIETTGKAVYITSTLPYVVLTIFLIRGLTLKGATNGIVFLFTPN 258

Qy    289 LVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLAT 348
          : ::  |  | :   |||   | |||:| |||||  |||  |  :|| || ||||
Db    259 VTELAQPDTWLDAGAQVFFSFSLAFGGLISFSSYNSVHNNCEKDSVIVSIINGFTSVYVA 318

Qy    349 LVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSR----DLIPPHVNFSHLTTKDYME 404
          :||: |:||:|    : |           :||:|:     ||  |  |   :| :::::
Db    319 IVVYSVIGFRATQRYDDC----------FSTNILTLINGFDL--PEGN---VTQENFVD 362

Qy    405 MDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFL 464
          |           :: |    | :   | ::|:|||||||| ||||:|  | || |||:||:
```

```
Db    363 MQQRCNASDPAAYAQLVFQTCDINAFLSEAVEGTGLAFIVFTEAITKMPLSPLWSVLFFI 422

Qy    465 MLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFV 520
          ||  ||| || | | |:  |: |      | |||: |   |: |||:| :|   || |::
Db    423 MLFCLGLSSMFGNMEGVVVPLQDLRVIPPKWPKEVLTGLICLGTFLIGFIFTLNSGQYWL 482

Qy    521 TMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCM 580
          :: | |: ::|| :|   |  :| ::||  :| :::  |:| :|  |:   |: |||| |
Db    483 SLLDSYAGSIPLLIIAFCEMFSVVYVYGVDRFNKDIEFMIGHKPNIFWQVTWRVVSPLLM 542

Qy    581 AVLTTASIIQLGVTPPAYSAW---IKEEAAERYLYFPNWPMALLITLIVVATLPIP--VV 635
           ::     :        || |    :|    : : :|||   ::: :  | :| ||   :
Db    543 LIIFLFFFVVEVSQELTYSIWDPGYEEFPKSQKISYPNWVYVVVVIVAGVPSLTIPGYAI 602

Qy    636 FVLRHFHLLSDGS-----NTLSVS 654
          : |   |    |      :||| :
Db    603 YKLIRNHCQKPGDHQGLVSTLSTA 626


RESULT 14
ABB59777
ID   ABB59777 standard; Protein; 662 AA.
XX
AC   ABB59777;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 6123.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL03880.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 6123; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   662 AA;

  Query Match             34.2%;  Score 1305;  DB 22;  Length 662;
  Best Local Similarity   40.7%;  Pred. No. 1.1e-124;
  Matches  248;  Conservative  109;  Mismatches  209;  Indels   44;  Gaps    4;

Qy     34 KQSVLNVAGEAGGKQKAVEEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
          |  :: :|  :  :|  || | : |:| :|:||: ::|| ||::|||||:|||||| |||
Db     28 KAELVAIATASADQQDIVEGEKEGEERESWDSKIMFLLATIGYAVGLGNVWRFPYLAQKN 87

Qy     94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
          ||||:|||| ::| | |||:|:||||:|||:|:|:|||| : | ||||| |  :|   |
Db     88 GGGAFLVPYFIMLCIQGIPIFYLELAIGQRLRKGAIGVWSQVSPYLGGIGISSAVVSYIV 147

Qy    154 GLYYNVIIGWSIFYFFKSFQYPLPWSECPV-VRNGSVAVVEAECEKSSATTYFWYREALD 212
          |||| || | : |   ||: ||||::||  :        | ||  || | ::|||  |
Db    148 ALYYNTIIAWCLIYLLHSFESPLPWADCPTRLYKNFTYDHEPECVASSPTQFYWYRTTLQ 207

Qy    213 ISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGL 272
          |:|:      |: | : |:| | :  : : :|| ||||::| :::|||||  |  ||:
Db    208 CSESVDMPENFNYHMAIALIVSWFLVYICMVQGITSSGKIVYMTAIFPYVVLIIFFFRGI 267

Qy    273 LLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFD 332
          |:|| ||: |:|||:   :||| || |  ||:|| ||| |||:| |||||  :|||: |
Db    268 TLKGAADGVAHLFTPRWETLLDPVVWLEAGTQIFFSLGLAFGGLIAFSSYNPANNNCYRD 327

Qy    333 GALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHV 392
          |||  |  ||: | :||| |:||||    ::|  |    :                .
Db    328 AILVSLTNCGTSMFAGVVVFSVIGFKATATFDRCTEERNGLVA---------------Q 371

Qy    393 NFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLLEDELDKSVQGTGLAFIAFTEAMTHF 452
          | :|                              |  | |: ||  |  |||||| ||||:  |
Db    372 NKTH----------------------NLPVCDLQTELANSASGTGLAFIIFTEAINQF 407

Qy    453 PTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIIDTS---KVPKEMFTVGCCVFTFLVGL 509
          | :  |:|:|||||  ||: |  ||: |: | ::|      :|||      |    : :
Db    408 PGAQLWAVLFFLMLFTLGIDSQFGTLEGVVTSLVDMKLFPNLPKEYIVGALCFSCCTISM 467

Qy    510 LFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRFYF 569
          |   :|:|   : | ::   || :| : | :::::||| ::|  ::  | | ||  ::
Db    468 CFANGAGSYIFQLMDSFAGNFPLLIIALFECLSISYIYGVRRFSDDIEMMTGSRPNFYWM 527

Qy    570 YMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWPMALLITLIVVAT 629
          : ||::||  |  :  ||  ||     :| |||  : |   : :|:| : :   ||: :
Db    528 FCWKYLSPCAMVTILLASFYQLLTEGSSYPAWIGSKGATEGMEWPHWCIVVAFFLILSSI 587

Qy    630 LPIPVVFVLR 639
          | ||:| |||
Db    588 LWIPIVAVLR 597


RESULT 15
AAW07635
ID   AAW07635 standard; Protein; 616 AA.
XX
AC   AAW07635;
XX
DT   05-FEB-1997  (first entry)
XX
DE   Rat brain derived orphan transporter, rB21a.
XX
KW   Rat; brain; orphan; transporter; rB21a; identification; drug;
KW   therapeutic agent; neurological; psychiatric; disorder;
KW   alleviation; abnormality; treatment; migraine headache; swelling;
KW   injury; hypoxia; seizure; stroke; probe; primer; detection;
KW   therapy; production; transgenic animal; antibody; inhibition.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   Domain          38..58
FT                   /note= "putative transmembrane domain I"
FT   Domain          65..85
FT                   /note= "putative transmembrane domain II"
FT   Domain          110..130
FT                   /note= "putative transmembrane domain III"
FT   Domain          194..214
FT                   /note= "putative transmembrane domain IV"
```

```
FT   Domain          219..239
FT                   /note= "putative transmembrane domain V"
FT   Domain          272..292
FT                   /note= "putative transmembrane domain VI"
FT   Domain          301..321
FT                   /note= "putative transmembrane domain VII"
FT   Domain          408..428
FT                   /note= "putative transmembrane domain VIII"
FT   Domain          456..476
FT                   /note= "putative transmembrane domain IX"
FT   Domain          490..510
FT                   /note= "putative transmembrane domain X"
FT   Domain          527..547
FT                   /note= "putative transmembrane domain XI"
FT   Domain          582..602
FT                   /note= "putative transmembrane domain XII"
XX
PN   US5559021-A.
XX
PD   24-SEP-1996.
XX
PF   09-NOV-1993;   93US-0149100.
XX
PR   09-NOV-1993;   93US-0149100.
XX
PA   (SYNA-) SYNAPTIC PHARM CORP.
XX
PI   Smith KE,  Weinshank RL;
XX
DR   WPI; 1996-442411/44.
DR   N-PSDB; AAT44120.
XX
PT   New isolated nucleic acid encoding orphan transporter, rB21a, from
PT   rat brain - homologous to neuro:transmitter transporters, used to
PT   develop prods. for detection or therapy, e.g. for migraine or brain
PT   swelling.
XX
PS   Claim 1; Columns 23-28; 30pp; English.
XX
CC   The present sequence is the rat brain derived orphan transporter,
CC   rB21a, which may be used to identify cpds. which specifically
CC   interact with and bind to it, for the development of therapeutic
CC   agents for neurological and psychiatric disorders. The drugs can
CC   be used to alleviate abnormalities resulting from the over or
CC   under expression of the transporter, and to treat conditions such
CC   as migraine headaches, and brain swelling after injury, hypoxia,
CC   seizures and stroke. The rB21a encoding cDNA can be used to produce
CC   the transporter, or oligonucleotide compsns. for the detection or
CC   therapy of rB21a related disease, or to produce transgenic animals.
CC   The transporter protein can also be used to produce antibodies for
CC   rB21a detection and inhibition. The rB21a cDNA was obtd. by
CC   screening a rat brain cDNA library with probes based on rat GABA
CC   transporter (GAT-1) cDNA.
XX
SQ   Sequence   616 AA;

  Query Match             34.1%;  Score 1302.5;  DB 17;  Length 616;
  Best Local Similarity   43.1%;  Pred. No. 1.8e-124;
  Matches  265;  Conservative  122;  Mismatches  191;  Indels   37;  Gaps   14;

Qy      38 LNVAGEAGGKQKAVEEELDAED----RPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKN 93
           |  :    |    |:    :|  | |     || | : ||:: | | ::|||||:||||||||
Db       3 LAIKRRASRGQRPGPDEKRARDMEKARPQWGNPLQFVFACISYAVGLGNVWRFPYLCQMY 62

Qy      94 GGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFV 153
           |||::|||||::||: |:|| :|||||||:|:|||| |  | | | |:| :| :|  |:
Db      63 GGGSFLVPYLIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFL 122

Qy     154 GLYYNVIIGWSIFYFFKSFQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDI 213
            :|||||  |  :| | ||| ||||| ||:  | :    : ||||:|:| |||||: |:|
Db     123 SMYYNVINAWGFWYLFHSFQDPLPWSVCPL--NSNRTGYDEECEKASSTQYFWYRKTLNI 180

Qy     214 SDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLL 273
           | || |:||: |:  ||| : | :  : : :| :|:|||:||::| || ||  :|||||
Db     181 SPSIQENGGVQWEPALCLTLAWLMVYLCILRGTESTGKVVYFTALMPYCVLIIYLVRGLT 240

Qy     274 LRGAVDGILHMFTPKLVKMLDPQVWREVATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDG 333
           | || :|:::|||||: :: :|: |   |||:|| |||||| :| |:|||:  |:|
Db     241 LHGATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNEPSNDCQKHA 300

Qy     334 ALVSFINFFTSVLATLVVFVVLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVN 393
            :|| ||  ||: |::| | : |||     | |:     |::  | ||  | ||
Db     301 VIVSVINSSTSIFASIVTFSIYGFKATFNYENCL----NKVI-LLLTN--SFDL-----E 348

Qy     394 FSHLTTKDYMEMDNVIMTVKEDQFSAL--GLDPCLLEDELDKSVQGTGLAFIAFTEAMTH 451
              ||  :  |: : : :   :::| :   :  | || ||: :|||||||||| : ||: :
Db     349 DGFLTASNLEEVKDYLASTYPNKYSEVFPHIRNCSLESELNTAVQGTGLAFIVYAEAIKN 408

Qy     452 FPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID----TSKVPKEMFTVGCCVFTFLV 507
              |  |||::| ||: ||:|||:|  | | ||: |    :| :|||  :   |:    |
Db     409 MEVSQLWSVLYFFMLLMLGMGSMLGNTAAILTPLTDSKVISSYLPKEAISGLVCLINCAV 468

Qy     508 GLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYGPKKFMQELTEMLGFRPYRF 567
           |::|   :|||:  :|:||:|||  | |||::| ||| ::|| ::|  :|  | | ||  :
Db     469 GMVFTMEAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLRAMTG-RPLNW 527

Qy     568 YF-YMWKFVSPLC---MAVLTTASIIQLGVTPPAYSAWIKEEAAERYLYFPNWP---MAL 620
           |:  || |||||    : :   :  |  |     | ||   :| :  |   ::|   :|:
Db     528 YWKAMWAFVSPLLIIGLFIFYLSDYILTGTL--QYQAW---DATQGQLVTKDYPPHALAV 582

Qy     621 LITLIVVATLPIPVV 635
           :  |:  :|: ||:|
Db     583 IGLLVASSTMCIPLV 597


Search completed: April 21, 2003, 13:55:26
Job time : 43 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:08:09 ; Search time 41 Seconds
(without alignments)
3653.567 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 727
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	3.0	241	11	Q8R0C2	Q8r0c2 mus musculu
2	19	2.6	249	4	Q9NW50	Q9nw50 homo sapien
3	16	2.2	265	6	Q9XS32	Q9xs32 bos taurus
4	16	2.2	289	4	Q9H9F5	Q9h9f5 homo sapien
5	16	2.2	289	4	Q8TBM6	Q8tbm6 homo sapien
6	15	2.1	744	5	Q9V8I8	Q9v8i8 drosophila
7	14	1.9	397	4	Q9BYZ7	Q9byz7 homo sapien
8	14	1.9	577	11	Q91XG6	Q91xg6 mus musculu
9	14	1.9	615	11	O88576	O88576 mus musculu
10	14	1.9	628	4	Q96N87	Q96n87 homo sapien
11	13	1.8	615	11	Q62687	Q62687 rattus norv
12	12	1.7	674	5	Q8WPM9	Q8wpm9 oikopleura
13	11	1.5	502	16	Q9CMU0	Q9cmu0 pasteurella
14	11	1.5	505	16	Q9K7C5	Q9k7c5 bacillus ha
15	11	1.5	511	16	Q9JXM3	Q9jxm3 neisseria m
16	11	1.5	511	16	Q9JWA5	Q9jwa5 neisseria m
17	11	1.5	614	5	Q9U2E3	Q9u2e3 caenorhabdi
18	10	1.4	102	5	Q9NF30	Q9nf30 caenorhabdi
19	10	1.4	527	16	Q9KQI6	Q9kqi6 vibrio chol
20	10	1.4	635	5	O18288	O18288 caenorhabdi
21	10	1.4	706	5	Q9XUG5	Q9xug5 caenorhabdi
22	10	1.4	728	5	Q9U2E0	Q9u2e0 caenorhabdi
23	9	1.2	432	17	Q8TXS1	Q8txs1 methanopyru
24	9	1.2	453	16	Q9KDT3	Q9kdt3 bacillus ha
25	9	1.2	489	17	Q8TQM2	Q8tqm2 methanosarc
26	9	1.2	518	13	Q90WD3	Q90wd3 oncorhynchu
27	9	1.2	592	4	Q9NP91	Q9np91 homo sapien
28	9	1.2	592	11	Q8VDB9	Q8vdb9 mus musculu
29	9	1.2	616	4	Q8TF10	Q8tf10 homo sapien
30	9	1.2	616	11	Q64093	Q64093 rattus norv
31	9	1.2	635	11	Q91WT6	Q91wt6 mus musculu
32	9	1.2	635	11	O88575	O88575 mus musculu
33	9	1.2	768	5	P90890	P90890 caenorhabdi
34	8	1.1	128	13	Q9DFY5	Q9dfy5 rana catesb
35	8	1.1	155	2	Q9Z5Y2	Q9z5y2 frankia sp.
36	8	1.1	174	8	Q9ME78	Q9me78 drosophila
37	8	1.1	174	8	Q9ME49	Q9me49 drosophila
38	8	1.1	174	8	Q9MD81	Q9md81 drosophila
39	8	1.1	174	8	Q9MGM5	Q9mgm5 drosophila
40	8	1.1	174	8	Q9MGL6	Q9mgl6 drosophila
41	8	1.1	264	16	Q9PM93	Q9pm93 campylobact
42	8	1.1	318	16	Q9KMZ4	Q9kmz4 vibrio chol
43	8	1.1	333	5	Q22352	Q22352 caenorhabdi
44	8	1.1	431	16	Q9RWL0	Q9rwl0 deinococcus
45	8	1.1	449	16	Q9CMM5	Q9cmm5 pasteurella

ALIGNMENTS

RESULT 1
Q8R0C2
ID Q8R0C2 PRELIMINARY; PRT; 241 AA.
AC Q8R0C2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 27.0 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC027078; AAH27078.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 241 AA; 27030 MW; AA15BDCBF5BBE07C CRC64;

Query Match 3.0%; Score 22; DB 11; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 511 FVQRSGNYFVTMFDDYSATLPL 532
||||||||||||||||||||||
Db 24 FVQRSGNYFVTMFDDYSATLPL 45

RESULT 2
Q9NW50
ID Q9NW50 PRELIMINARY; PRT; 249 AA.
AC Q9NW50;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA FLJ10316 fis, clone NT2RM2000422, highly similar to sodium- and

```
DE   chloride-dependent transporter NTT73.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK001178; BAA91536.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   249 AA;  27873 MW;  64395F571A377377 CRC64;

  Query Match             2.6%;  Score 19;  DB 4;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    435 VQGTGLAFIAFTEAMTHFP 453
          |||||||||||||||||||
Db    152 VQGTGLAFIAFTEAMTHFP 170


RESULT 3
Q9XS32
ID   Q9XS32      PRELIMINARY;      PRT;   265 AA.
AC   Q9XS32;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Orphan transporter short splicing variant.
GN   BV7-3S3 OR BV7-3S2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sakata K., Shimada S.;
RT   "Orphan transporter.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB020857; BAA77226.1; -.
DR   EMBL; AB020855; BAA77224.1; -.
DR   EMBL; AB020856; BAA77225.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   265 AA;  29687 MW;  20F95214D2AD26F9 CRC64;

  Query Match             2.2%;  Score 16;  DB 6;  Length 265;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     84 WRFPYLCQKNGGGAYL 99
          ||||||||||||||||
Db     85 WRFPYLCQKNGGGAYL 100


RESULT 4
Q9H9F5
ID   Q9H9F5      PRELIMINARY;      PRT;   289 AA.
AC   Q9H9F5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ12791 fis, clone NT2RP2001991, highly similar to sodium- and
DE   chloride-dependent transporter NTT73.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022853; BAB14274.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   289 AA;  32264 MW;  AB0768CD8F74BF47 CRC64;

  Query Match             2.2%;  Score 16;  DB 4;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     84 WRFPYLCQKNGGGAYL 99
          ||||||||||||||||
Db     85 WRFPYLCQKNGGGAYL 100


RESULT 5
Q8TBM6
ID   Q8TBM6      PRELIMINARY;      PRT;   289 AA.
AC   Q8TBM6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 32.2 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC022253; AAH22253.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   289 AA;  32248 MW;  AB0768DB5574BF47 CRC64;

  Query Match             2.2%;  Score 16;  DB 4;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     84 WRFPYLCQKNGGGAYL 99
          ||||||||||||||||
Db     85 WRFPYLCQKNGGGAYL 100


RESULT 6
Q9V8I8
ID   Q9V8I8      PRELIMINARY;      PRT;   744 AA.
AC   Q9V8I8;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CG5226 protein.
GN   CG5226.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003799; AAF57678.1; -.
DR   FlyBase; FBgn0034355; CG5226.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   744 AA;  82953 MW;  ED4A742B3E0E483B CRC64;

  Query Match             2.1%;  Score 15;  DB 5;  Length 744;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      77 SVGLGNIWRFPYLCQ 91
           |||||||||||||||
Db      72 SVGLGNIWRFPYLCQ 86


RESULT 7
Q9BYZ7
ID   Q9BYZ7      PRELIMINARY;      PRT;   397 AA.
AC   Q9BYZ7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium channel-like protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Londono-Vallejo J.A.;
RT   "Sequence of a BAC carrying the entire hTERT gene.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY007685; AAG23290.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   397 AA;  43784 MW;  E70588E4D7FD9CF6 CRC64;

  Query Match             1.9%;  Score 14;  DB 4;  Length 397;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLCQ 91
           ||||||||||||||
Db      36 VGLGNIWRFPYLCQ 49


RESULT 8
Q91XG6
ID   Q91XG6      PRELIMINARY;      PRT;   577 AA.
AC   Q91XG6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to X transporter protein 2.
GN   XTRP2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC010748; AAH10748.1; -.
DR   MGD; MGI:1336892; Xtrp2.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   577 AA;  65254 MW;  5F84F7C5CF1B14E3 CRC64;

  Query Match             1.9%;  Score 14;  DB 11;  Length 577;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLCQ 91
           ||||||||||||||
Db      35 VGLGNIWRFPYLCQ 48


RESULT 9
O88576
ID   O88576      PRELIMINARY;      PRT;   615 AA.
AC   O88576; O88577; O88578; O88579; O88580; O88581;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
```

```
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium- and chloride-dependent transporter XTRP2.
GN   XTRP2 OR XT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC   TISSUE=KIDNEY;
RX   MEDLINE=99131080; PubMed=9932288;
RA   Nash S.R., Giros B., Kingsmore S.F., Kim K.M., El-Mestikawy S.,
RA   Dong Q., Fumagalli F., Seldin M.F., Caron M.G.;
RT   "Cloning, gene structure, and genomic localization of an orphan
RT   transporter from mouse kidney with six alternatively-spliced
RT   isoforms.";
RL   Recept. Channels 6:113-128(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: 6 ISOFORMS; A12 (SHOWN HERE), A11, B11, A10,
CC       B9 AND A8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
DR   EMBL; AF075262; AAC27757.1; -.
DR   EMBL; AF075263; AAC27758.1; -.
DR   EMBL; AF075264; AAC27759.1; -.
DR   EMBL; AF075265; AAC27760.1; -.
DR   EMBL; AF075266; AAC27761.1; -.
DR   EMBL; AF075267; AAC27762.1; -.
DR   MGD; MGI:1336892; Xtrp2.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; FALSE_NEG.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Alternative splicing.
FT   DOMAIN       1     26       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    27     47       1 (POTENTIAL).
FT   DOMAIN      48     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    53     73       2 (POTENTIAL).
FT   DOMAIN      74     83       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    84    104       3 (POTENTIAL).
FT   DOMAIN     105    105       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   106    126       4 (POTENTIAL).
FT   DOMAIN     127    177       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   178    198       5 (POTENTIAL).
FT   DOMAIN     199    206       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   207    227       6 (POTENTIAL).
FT   DOMAIN     228    288       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   289    309       7 (POTENTIAL).
FT   DOMAIN     310    397       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   398    418       8 (POTENTIAL).
FT   DOMAIN     419    441       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   442    462       9 (POTENTIAL).
FT   DOMAIN     463    472       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   473    493       10 (POTENTIAL).
FT   DOMAIN     494    520       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   521    541       11 (POTENTIAL).
FT   DOMAIN     542    570       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   571    591       12 (POTENTIAL).
FT   DOMAIN     592    615       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD   353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   377    422       SASGPGLAFIVFTEAVLHMPGASVWSVLFFGMLFTLGLSSM
FT                               FGNME -> VLLCGLCSSLGCCLPWVCPPCLGTWRVSLHHY
FT                               WTWGSYPKVYPRRS (IN ISOFORM A8).
FT   VARSPLIC   377    444       SASGPGLAFIVFTEAVLHMPGASVWSVLFFGMLFTLGLSSM
FT                               FGNMEGVITPLLDMGILPKGIPKEVMT -> PTWKQISGAR
FT                               VLGEGCARLTSRVCEASVLP (IN ISOFORM B11 AND
FT                               ISOFORM B9).
FT   VARSPLIC   445    454       MISSING (IN ISOFORM A11).
FT   VARSPLIC   499    552       FCDDIEWMTGRRPGLYWQVTWRVVSPMLLFGIFLSYIVLLI
FT                               QTPPSYKAWNPQY -> NIFPQERRSSTQAGCRSPVCSCPS
FT                               CPHCGSLELLWLSYCPSTNRGGRLRIWKVV (IN
FT                               ISOFORM A10 AND ISOFORM B9).
FT   VARSPLIC   553    615       MISSING (IN ISOFORM A10).
SQ   SEQUENCE   615 AA;  69228 MW;  A7925C0397FC63F8 CRC64;

  Query Match             1.9%;  Score 14;  DB 11;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLCQ 91
           ||||||||||||||
Db      35 VGLGNIWRFPYLCQ 48


RESULT 10
Q96N87
ID   Q96N87      PRELIMINARY;      PRT;   628 AA.
AC   Q96N87;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ31236 fis, clone KIDNE2004828, moderately similar to Mus
DE   musculus orphan transporter isoform A12 (Xtrp2) mRNA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA   Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA   Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA   Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK055798; BAB71018.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   628 AA;  70924 MW;  90EA71CD12C6DC9D CRC64;

  Query Match             1.9%;  Score 14;  DB 4;  Length 628;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLCQ 91
           ||||||||||||||
Db      36 VGLGNIWRFPYLCQ 49


RESULT 11
Q62687
ID   Q62687      PRELIMINARY;      PRT;   615 AA.
AC   Q62687;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Renal osmotic stress-induced Na-Cl organic solute cotransporter.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY CORTEX;
```

```
RX   MEDLINE=95029937; PubMed=7943364;
RA   Wasserman J.C., Delpire E., Tonidandel W., Kojima R., Gullans S.R.;
RT   "Molecular characterization of ROSIT, a renal osmotic stress-induced
RT   Na(+)-Cl(-)-organic solute cotransporter.";
RL   Am. J. Physiol. 267:F688-F694(1994).
DR   EMBL; U12973; AAC13771.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   615 AA;  69556 MW;  F41E7DE1DC276918 CRC64;

  Query Match             1.8%;  Score 13;  DB 11;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLC 90
           |||||||||||||
Db      35 VGLGNIWRFPYLC 47


RESULT 12
Q8WPM9
ID   Q8WPM9      PRELIMINARY;      PRT;   674 AA.
AC   Q8WPM9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to glycine transporter.
GN   BAC001.6.
OS   Oikopleura dioica.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC   Oikopleuridae; Oikopleura.
OX   NCBI_TaxID=34765;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Seo H.C., Kube M., Edvardsen R.B., Jensen M.F., Beck A., Spriet E.,
RA   Gorsky G., Thompson E.M., Lehrach H., Reinhardt R., Chourrout D.;
RT   "The marine chordate Oikopleura dioica has a miniature genome.";
RL   Science 0:0-0(2001).
DR   EMBL; AF374376; AAL56437.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; UNKNOWN_1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
SQ   SEQUENCE   674 AA;  75155 MW;  9D318B68D64416E6 CRC64;

  Query Match             1.7%;  Score 12;  DB 5;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 0.0024;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYL 89
           ||||||||||||
Db      97 VGLGNIWRFPYL 108


RESULT 13
Q9CMU0
ID   Q9CMU0      PRELIMINARY;      PRT;   502 AA.
AC   Q9CMU0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein PM0718.
GN   PM0718.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006109; AAK02802.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   502 AA;  53982 MW;  06ADEE74DECF1107 CRC64;

  Query Match             1.5%;  Score 11;  DB 16;  Length 502;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||||
Db      26 VGLGNIWRFPY 36


RESULT 14
Q9K7C5
ID   Q9K7C5      PRELIMINARY;      PRT;   505 AA.
AC   Q9K7C5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium-dependent transporter.
GN   BH3439.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001518; BAB07158.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Complete proteome.
SQ   SEQUENCE   505 AA;  54296 MW;  607C0F6A00D67B07 CRC64;

  Query Match             1.5%;  Score 11;  DB 16;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||||
Db      20 VGLGNIWRFPY 30


RESULT 15
Q9JXM3
ID   Q9JXM3      PRELIMINARY;      PRT;   511 AA.
AC   Q9JXM3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
```

```
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sodium-and chloride-dependent transporter.
GN   NMB1975.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002546; AAF42303.1; -.
DR   TIGR; NMB1975; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Complete proteome.
SQ   SEQUENCE   511 AA;  54905 MW;  270B4E731EE7A1DD CRC64;

  Query Match             1.5%;  Score 11;  DB 16;  Length 511;
  Best Local Similarity   100.0%;  Pred. No. 0.022;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||| 
Db      27 VGLGNIWRFPY 37


Search completed: April 21, 2003, 14:11:06
Job time : 45 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:07:19 ; Search time 14 Seconds
(without alignments)
2153.806 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 727
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	11.4	727	1	NTT4_RAT	P31662 rattus norv
2	52	7.2	439	1	NTT4_HUMAN	Q9h1v8 homo sapien
3	51	7.0	225	1	NTT4_BOVIN	Q28001 bos taurus
4	22	3.0	729	1	NTT7_BOVIN	Q9xs59 bos taurus
5	22	3.0	729	1	NTT7_RAT	Q08469 rattus norv
6	22	3.0	730	1	NTT7_HUMAN	Q9h2j7 homo sapien
7	11	1.5	508	1	Y736_HAEIN	P44849 haemophilus
8	11	1.5	615	1	NTDO_CAEEL	Q03614 caenorhabdi
9	9	1.2	518	1	CP3R_ONCMY	O42563 oncorhynchu
10	9	1.2	630	1	S6A4_CAVPO	O35899 cavia porce
11	9	1.2	630	1	S6A4_MOUSE	Q60857 mus musculu
12	9	1.2	630	1	S6A4_RAT	P31652 rattus norv
13	9	1.2	636	1	S6A7_HUMAN	Q99884 homo sapien
14	9	1.2	637	1	S6A7_RAT	P28573 rattus norv
15	8	1.1	174	1	NU6M_DROYA	P07709 drosophila
16	8	1.1	492	1	YD19_METJA	Q58715 methanococc
17	8	1.1	611	1	YCR3_YEAST	P25351 saccharomyc
18	8	1.1	690	1	ATKB_SYNY3	P73867 synechocyst
19	8	1.1	1278	1	NPC1_HUMAN	O15118 homo sapien
20	7	1.0	73	1	MALE_PHOLU	P41130 photorhabdu
21	7	1.0	107	1	YE11_YEAST	P40097 saccharomyc
22	7	1.0	112	1	UCN2_MOUSE	Q99m18 mus musculu
23	7	1.0	135	1	H32_ORYSA	P08860 oryza sativ
24	7	1.0	158	1	UREE_KLEAE	P18317 klebsiella
25	7	1.0	174	1	NU6M_DROME	P18933 drosophila
26	7	1.0	174	1	NU6M_PAPHA	Q9zxx5 papio hamad
27	7	1.0	215	1	ERD2_ARATH	P35402 arabidopsis
28	7	1.0	231	1	YOHK_ECOLI	P33373 escherichia
29	7	1.0	250	1	PSA7_DICDI	P34120 dictyosteli
30	7	1.0	252	1	UNG_HSVSA	Q01019 herpesvirus
31	7	1.0	279	1	EUTJ_SALTY	P41794 salmonella
32	7	1.0	314	1	RLPA_HELPJ	Q9zj38 helicobacte
33	7	1.0	315	1	RLPA_HELPY	O26091 helicobacte
34	7	1.0	317	1	HEMZ_CHLMU	Q9pjq6 chlamydia m
35	7	1.0	317	1	MSHR_CAPHI	P56444 capra hircu
36	7	1.0	317	1	MSHR_OVIMO	P56447 ovibos mosc
37	7	1.0	317	1	MSHR_SHEEP	O19037 ovis aries
38	7	1.0	328	1	YCDU_ECOLI	P75910 escherichia
39	7	1.0	335	1	PAPG_ECOLI	P13720 escherichia
40	7	1.0	338	1	OTCA_BUCAI	P57449 buchnera ap
41	7	1.0	347	1	GABC_HUMAN	Q06545 homo sapien
42	7	1.0	347	1	GABC_MOUSE	Q00421 mus musculu
43	7	1.0	347	1	VP39_NPVAC	P17499 autographa
44	7	1.0	351	1	EGSA_SULSO	Q9uxe7 sulfolobus
45	7	1.0	367	1	CTNS_MOUSE	P57757 mus musculu

ALIGNMENTS

RESULT 1
NTT4_RAT
ID NTT4_RAT STANDARD; PRT; 727 AA.
AC P31662;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Orphan sodium- and chloride-dependent neurotransmitter transporter
DE NTT4.
GN NTT4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93114444; PubMed=8093354;
RA Liu Q.-R., Mandiyan S., Lopez-Corcuera B., Nelson H., Nelson N.;
RT "A rat brain cDNA encoding the neurotransmitter transporter with an
RT unusual structure.";
RL FEBS Lett. 315:114-118(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94125086; PubMed=8294906;
RA el Mestikawy S., Giros B., Pohl M., Hamon M., Kingsmore S.F.,
RA Seldin M.F., Caron M.G.;
RT "Characterization of an atypical member of the Na+/Cl(-)-dependent
RT transporter family: chromosomal localization and distribution in
RT GABAergic and glutamatergic neurons in the rat brain.";
RL J. Neurochem. 62:445-455(1994).
CC -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CENTRAL NERVOUS
CC SYSTEM AND IS MORE ABUNDANT IN THE CEREBELLUM AND THE CEREBRAL
CC CORTEX.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNF).
CC ----------

CC ----------
DR EMBL; L06434; AAB24776.1; -.
DR EMBL; S68944; AAC60673.1; -.
DR PIR; S27043; S27043.
DR InterPro; IPR000175; Na/ntran_symport.
DR Pfam; PF00209; SNF; 1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/ntran_symport; 2.
DR PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.

```
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     68       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     69     89       1 (POTENTIAL).
FT   TRANSMEM     97    116       2 (POTENTIAL).
FT   TRANSMEM    141    161       3 (POTENTIAL).
FT   DOMAIN      162    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    243       4 (POTENTIAL).
FT   TRANSMEM    252    269       5 (POTENTIAL).
FT   TRANSMEM    305    322       6 (POTENTIAL).
FT   TRANSMEM    334    355       7 (POTENTIAL).
FT   DOMAIN      356    451       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    452    471       8 (POTENTIAL).
FT   TRANSMEM    495    513       9 (POTENTIAL).
FT   TRANSMEM    529    549       10 (POTENTIAL).
FT   TRANSMEM    570    591       11 (POTENTIAL).
FT   TRANSMEM    619    641       12 (POTENTIAL).
FT   DOMAIN      642    727       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    393    393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    261    261       Y -> C (IN REF. 2).
FT   CONFLICT    499    499       G -> S (IN REF. 2).
SQ   SEQUENCE   727 AA;  81055 MW;  C676048C0A6BDF7C CRC64;

  Query Match             11.4%;  Score 83;  DB 1;  Length 727;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-76;
  Matches   83;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111

Qy     112 PLFFLELAVGQRIRRGSIGVWHY 134
           |||||||||||||||||||||||
Db     112 PLFFLELAVGQRIRRGSIGVWHY 134


RESULT 2
NTT4_HUMAN
ID   NTT4_HUMAN     STANDARD;      PRT;   439 AA.
AC   Q9H1V8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT4 (Fragment).
GN   NTT4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Clark G.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL137790; CAC19682.1; ALT_INIT.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; PARTIAL.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; PARTIAL.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   NON_TER       1      1
FT   TRANSMEM     17     34       6 (POTENTIAL).
FT   TRANSMEM     46     67       7 (POTENTIAL).
FT   DOMAIN       68    163       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    164    183       8 (POTENTIAL).
FT   TRANSMEM    207    225       9 (POTENTIAL).
FT   TRANSMEM    241    261       10 (POTENTIAL).
FT   TRANSMEM    282    303       11 (POTENTIAL).
FT   TRANSMEM    331    353       12 (POTENTIAL).
FT   DOMAIN      354    439       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   439 AA;  49011 MW;  44634308C281C740 CRC64;

  Query Match             7.2%;  Score 52;  DB 1;  Length 439;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-44;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     354 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEM 405
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEM 117


RESULT 3
NTT4_BOVIN
ID   NTT4_BOVIN     STANDARD;      PRT;   225 AA.
AC   Q28001;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT4 (Fragment).
GN   NTT4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RA   Jones E.M.C.;
RT   "Sodium-and chloride-dependent neurotransmitter transporters in bovine
RT   retina: identification and localization by in situ hybridization
RT   histochemistry.";
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U19593; AAA61578.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
```

```
FT   NON_TER       1      1
FT   TRANSMEM     <1      8       1 (POTENTIAL).
FT   TRANSMEM     16     35       2 (POTENTIAL).
FT   TRANSMEM     60     80       3 (POTENTIAL).
FT   DOMAIN       81    143       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    144    162       4 (POTENTIAL).
FT   TRANSMEM    171    188       5 (POTENTIAL).
FT   TRANSMEM    224   >225       6 (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25374 MW;  3FA15D33F76D15C3 CRC64;

  Query Match             7.0%;  Score 51;  DB 1;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-44;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      84 WRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHY 134
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 WRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAVGQRIRRGSIGVWHY 53


RESULT 4
NTT7_BOVIN
ID   NTT7_BOVIN     STANDARD;      PRT;   729 AA.
AC   Q9XS59;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT73 (Orphan transporter v7-3).
GN   NTT73.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RA   Sakata K.;
RT   "Orphan transporter.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB020854; BAA77223.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     90       1 (POTENTIAL).
FT   TRANSMEM     98    117       2 (POTENTIAL).
FT   TRANSMEM    142    162       3 (POTENTIAL).
FT   DOMAIN      163    225       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    226    244       4 (POTENTIAL).
FT   TRANSMEM    253    270       5 (POTENTIAL).
FT   TRANSMEM    306    323       6 (POTENTIAL).
FT   TRANSMEM    335    356       7 (POTENTIAL).
FT   DOMAIN      357    452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    453    472       8 (POTENTIAL).
FT   TRANSMEM    496    514       9 (POTENTIAL).
FT   TRANSMEM    530    550       10 (POTENTIAL).
FT   TRANSMEM    571    592       11 (POTENTIAL).
FT   TRANSMEM    620    642       12 (POTENTIAL).
FT   DOMAIN      643    729       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    213    213       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    383    383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    394    394       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   729 AA;  81681 MW;  332FD83349C196A9 CRC64;

  Query Match             3.0%;  Score 22;  DB 1;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533


RESULT 5
NTT7_RAT
ID   NTT7_RAT       STANDARD;      PRT;   729 AA.
AC   Q08469; Q63838;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT73 (Orphan transporter v7-3).
GN   NTT73.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93180651; PubMed=1363329;
RA   Uhl G.R., Kitayama S., Gregor P., Nanthakumar E., Persico A.M.,
RA   Shimada S.;
RT   "Neurotransmitter transporter family cDNAs in a rat midbrain library:
RT   'orphan transporters' suggest sizable structural variations.";
RL   Brain Res. Mol. Brain Res. 16:353-359(1992).
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L22022; AAA41729.1; -.
DR   EMBL; S56968; AAB25532.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     90       1 (POTENTIAL).
FT   TRANSMEM     98    117       2 (POTENTIAL).
```

```
FT   TRANSMEM     142     162       3 (POTENTIAL).
FT   DOMAIN       163     225       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     226     244       4 (POTENTIAL).
FT   TRANSMEM     253     270       5 (POTENTIAL).
FT   TRANSMEM     306     323       6 (POTENTIAL).
FT   TRANSMEM     335     356       7 (POTENTIAL).
FT   DOMAIN       357     452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     453     472       8 (POTENTIAL).
FT   TRANSMEM     496     514       9 (POTENTIAL).
FT   TRANSMEM     530     550       10 (POTENTIAL).
FT   TRANSMEM     571     592       11 (POTENTIAL).
FT   TRANSMEM     620     642       12 (POTENTIAL).
FT   DOMAIN       643     729       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     187     187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     383     383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     394     394       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT       1       3       MPK -> IP (IN REF. 1; AAB25532).
SQ   SEQUENCE   729 AA;  81596 MW;  ECDDC6403741B53E CRC64;

  Query Match             3.0%;  Score 22;  DB 1;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533


RESULT 6
NTT7_HUMAN
ID   NTT7_HUMAN     STANDARD;      PRT;   730 AA.
AC   Q9H2J7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orphan sodium- and chloride-dependent neurotransmitter transporter
DE   NTT73 (Orphan transporter v7-3).
GN   NTT73.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cerebellum;
RX   MEDLINE=20564208; PubMed=11112352;
RA   Farmer M.K., Robbins M.J., Medhurst A.D., Campbell D.A., Ellington K.,
RA   Duckworth M., Brown A.M., Middlemiss D.N., Price G.W., Pangalos M.N.;
RT   "Cloning and characterization of human NTT5 and v7-3: two orphan
RT   transporters of the Na(+)/Cl(-)-dependent neurotransmitter
RT   transporter gene family.";
RL   Genomics 70:241-252(2000).
CC   -!- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF265577; AAG41361.1; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN         1      69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      70      90       1 (POTENTIAL).
FT   TRANSMEM      98     117       2 (POTENTIAL).
FT   TRANSMEM     142     162       3 (POTENTIAL).
FT   DOMAIN       163     225       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     226     244       4 (POTENTIAL).
FT   TRANSMEM     253     270       5 (POTENTIAL).
FT   TRANSMEM     306     323       6 (POTENTIAL).
FT   TRANSMEM     335     356       7 (POTENTIAL).
FT   DOMAIN       357     452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     453     472       8 (POTENTIAL).
FT   TRANSMEM     496     514       9 (POTENTIAL).
FT   TRANSMEM     530     550       10 (POTENTIAL).
FT   TRANSMEM     571     592       11 (POTENTIAL).
FT   TRANSMEM     620     642       12 (POTENTIAL).
FT   DOMAIN       643     730       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     187     187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     213     213       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     383     383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     394     394       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   730 AA;  81836 MW;  45963118E06CFFE6 CRC64;

  Query Match             3.0%;  Score 22;  DB 1;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533


RESULT 7
Y736_HAEIN
ID   Y736_HAEIN     STANDARD;      PRT;   508 AA.
AC   P44849;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical sodium-dependent transporter HI0736.
GN   HI0736.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: PUTATIVE SODIUM-DEPENDENT TRANSPORTER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U32757; AAC22395.1; -.
DR   TIGR; HI0736; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; FALSE_NEG.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Hypothetical protein; Transport; Transmembrane; Symport;
KW   Complete proteome.
FT   TRANSMEM     19     39       POTENTIAL.
FT   TRANSMEM     47     67       POTENTIAL.
FT   TRANSMEM     92    112       POTENTIAL.
FT   TRANSMEM    153    173       POTENTIAL.
FT   TRANSMEM    186    206       POTENTIAL.
FT   TRANSMEM    228    248       POTENTIAL.
FT   TRANSMEM    272    292       POTENTIAL.
FT   TRANSMEM    325    345       POTENTIAL.
FT   TRANSMEM    365    385       POTENTIAL.
FT   TRANSMEM    394    414       POTENTIAL.
FT   TRANSMEM    432    452       POTENTIAL.
FT   TRANSMEM    469    489       POTENTIAL.
SQ   SEQUENCE   508 AA;  55177 MW;  5CC50526DE854BB9 CRC64;

  Query Match             1.5%;  Score 11;  DB 1;  Length 508;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||| 
Db      29 VGLGNIWRFPY 39


RESULT 8
NTDO_CAEEL
ID   NTDO_CAEEL     STANDARD;      PRT;   615 AA.
AC   Q03614; Q9XTK0;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Sodium-dependent dopamine transporter (DA transporter) (DAT).
GN   T23G5.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98440631; PubMed=9765501;
RA   Jayanthi L.D., Apparsundaram S., Malone M.D., Ward E., Miller D.M.,
RA   Eppler M., Blakely R.D.;
RT   "The Caenorhabditis elegans gene T23G5.5 encodes an antidepressant-and
RT   cocaine-sensitive dopamine transporter.";
RL   Mol. Pharmacol. 54:601-609(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [3]
RP   REVISIONS.
RA   Durbin R.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: TERMINATES THE ACTION OF DOPAMINE BY ITS HIGH
CC       AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
CC       STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF079899; AAC83661.1; -.
DR   EMBL; Z19158; CAA79575.2; -.
DR   EMBL; Z19156; CAA79575.2; JOINED.
DR   EMBL; Z19156; CAA79564.2; -.
DR   EMBL; Z19158; CAA79564.2; JOINED.
DR   PIR; S28306; S28306.
DR   WormPep; T23G5.5; CE25124.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 2.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     48       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     49     69       1 (POTENTIAL).
FT   TRANSMEM     74     94       2 (POTENTIAL).
FT   TRANSMEM    125    145       3 (POTENTIAL).
FT   DOMAIN      146    232       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    233    253       4 (POTENTIAL).
FT   TRANSMEM    258    278       5 (POTENTIAL).
FT   TRANSMEM    344    364       6 (POTENTIAL).
FT   TRANSMEM    371    391       7 (POTENTIAL).
FT   TRANSMEM    393    413       8 (POTENTIAL).
FT   TRANSMEM    440    460       9 (POTENTIAL).
FT   TRANSMEM    462    482       10 (POTENTIAL).
FT   TRANSMEM    515    535       11 (POTENTIAL).
FT   TRANSMEM    550    570       12 (POTENTIAL).
FT   DOMAIN      571    615       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    162    162       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   615 AA;  69265 MW;  0DA2876EB3AC8049 CRC64;

  Query Match             1.5%;  Score 11;  DB 1;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 0.0097;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPYLC 90
           ||||||||||| 
Db      60 LGNIWRFPYLC 70


RESULT 9
CP3R_ONCMY
ID   CP3R_ONCMY     STANDARD;      PRT;   518 AA.
AC   O42563;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Cytochrome P450 3A27 (EC 1.14.14.1) (CYPIIIA27).
```

```
GN   CYP3A27.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Shasta; TISSUE=Liver;
RX   MEDLINE=99045386; PubMed=9826429;
RA   Lee S.-J., Wang-Buhler J.-L., Cok I., Yu T.S., Yang Y.H.,
RA   Miranda C.L., Lech J., Buhler D.R.;
RT   "Cloning, sequencing, and tissue expression of CYP3A27, a new member
RT   of the CYP3A subfamily from embryonic and adult rainbow trout
RT   livers.";
RL   Arch. Biochem. Biophys. 360:53-61(1998).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC       oxidized flavoprotein + H(2)O.
CC   -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC   -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER
CC       TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,
CC       AND CARCINOGENS.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U96077; AAB82422.1; -.
DR   HSSP; P14779; 1JPZ.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum.
FT   BINDING    447    447       HEME (BY SIMILARITY).
SQ   SEQUENCE   518 AA;  59210 MW;  9B93AA12E617D0DF CRC64;

  Query Match             1.2%;  Score 9;  DB 1;  Length 518;
  Best Local Similarity   100.0%;  Pred. No. 0.91;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     619 ALLITLIVV 627
           |||||||||
Db      17 ALLITLIVV 25


RESULT 10
S6A4_CAVPO
ID   S6A4_CAVPO     STANDARD;      PRT;   630 AA.
AC   O35899;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium-dependent serotonin transporter (5HT transporter) (5HTT).
GN   SLC6A4.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestinal mucosa;
RX   MEDLINE=96180949; PubMed=8601815;
RA   Wade P.R., Chen J., Jaffe B., Kassem I.S., Blakely R.D., Gershon M.D.;
RT   "Localization and function of a 5-HT transporter in crypt epithelia
RT   of the gastrointestinal tract.";
RL   J. Neurosci. 16:2352-2364(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestinal mucosa;
RA   Chen J., Wade P.R., Rothman T.P., Gershon M.D.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: TERMINATES THE ACTION OF SEROTONINE BY ITS HIGH
CC       AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
CC       STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U84498; AAB82737.1; -.
DR   InterPro; IPR002437; 5HT_transporter.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   Pfam; PF03491; 5HT_transporter; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    108       1 (POTENTIAL).
FT   TRANSMEM    116    135       2 (POTENTIAL).
FT   TRANSMEM    160    180       3 (POTENTIAL).
FT   DOMAIN      181    252       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    253    271       4 (POTENTIAL).
FT   TRANSMEM    280    297       5 (POTENTIAL).
FT   TRANSMEM    333    350       6 (POTENTIAL).
FT   TRANSMEM    362    383       7 (POTENTIAL).
FT   TRANSMEM    417    436       8 (POTENTIAL).
FT   TRANSMEM    464    482       9 (POTENTIAL).
FT   TRANSMEM    498    518       10 (POTENTIAL).
FT   TRANSMEM    539    558       11 (POTENTIAL).
FT   TRANSMEM    577    595       12 (POTENTIAL).
FT   DOMAIN      596    630       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    217    217       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   630 AA;  70113 MW;  EE2A717DD9926D2F CRC64;

  Query Match             1.2%;  Score 9;  DB 1;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      99 LGNIWRFPY 107


RESULT 11
S6A4_MOUSE
ID   S6A4_MOUSE     STANDARD;      PRT;   630 AA.
AC   Q60857; O35241;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
```

```
DE   Sodium-dependent serotonin transporter (5HT transporter) (5HTT).
GN   SLC6A4 OR HTT OR SET.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97189261; PubMed=9037532;
RA   Chang A.S., Chang S.M., Starnes D.M., Schroeter S., Bauman A.L.,
RA   Blakely R.D.;
RT   "Cloning and expression of the mouse serotonin transporter.";
RL   Brain Res. Mol. Brain Res. 43:185-192(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129;
RX   MEDLINE=97225900; PubMed=9073170;
RA   Bengel D., Heils A., Petri S., Seemann M., Glatz K., Andrews A.,
RA   Murphy D.L., Lesch K.P.;
RT   "Gene structure and 5'-flanking regulatory region of the murine
RT   serotonin transporter.";
RL   Brain Res. Mol. Brain Res. 44:286-292(1997).
RN   [3]
RP   SEQUENCE OF 1-114 FROM N.A.
RC   STRAIN=BALB/c;
RA   Saito N., Sakai N., Kobayashi S., Fujimoto M., Morikawa O.,
RA   Ikegaki N.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 69-630 FROM N.A.
RX   MEDLINE=93283858; PubMed=8507984;
RA   Gregor P., Patel A., Shimada S., Lin C.L., Rochelle J.M., Kitayama S.,
RA   Seldin M.F., Uhl G.R.;
RT   "Murine serotonin transporter: sequence and localization to
RT   chromosome 11.";
RL   Mamm. Genome 4:283-284(1993).
CC   -!- FUNCTION: TERMINATES THE ACTION OF SEROTONINE BY ITS HIGH
CC       AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
CC       STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   -!- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT IN POSITION 195.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF013604; AAB67172.1; -.
DR   EMBL; Y08870; CAA70092.1; -.
DR   EMBL; Y08871; CAA70092.1; JOINED.
DR   EMBL; Y08872; CAA70092.1; JOINED.
DR   EMBL; Y08873; CAA70092.1; JOINED.
DR   EMBL; Y08874; CAA70092.1; JOINED.
DR   EMBL; Y08875; CAA70092.1; JOINED.
DR   EMBL; Y08876; CAA70092.1; JOINED.
DR   EMBL; Y08877; CAA70092.1; JOINED.
DR   EMBL; Y08878; CAA70092.1; JOINED.
DR   EMBL; Y08879; CAA70092.1; JOINED.
DR   EMBL; Y08880; CAA70092.1; JOINED.
DR   EMBL; U26452; AAA84750.1; -.
DR   EMBL; X66119; -; NOT_ANNOTATED_CDS.
DR   MGD; MGI:96285; Slc6a4.
DR   InterPro; IPR002437; 5HT_transporter.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   Pfam; PF03491; 5HT_transporter; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    108       1 (POTENTIAL).
FT   TRANSMEM    116    135       2 (POTENTIAL).
FT   TRANSMEM    160    180       3 (POTENTIAL).
FT   DOMAIN      181    252       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    253    271       4 (POTENTIAL).
FT   TRANSMEM    280    297       5 (POTENTIAL).
FT   TRANSMEM    333    350       6 (POTENTIAL).
FT   TRANSMEM    362    383       7 (POTENTIAL).
FT   TRANSMEM    417    436       8 (POTENTIAL).
FT   TRANSMEM    464    482       9 (POTENTIAL).
FT   TRANSMEM    498    518       10 (POTENTIAL).
FT   TRANSMEM    539    558       11 (POTENTIAL).
FT   TRANSMEM    577    595       12 (POTENTIAL).
FT   DOMAIN      596    630       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    217    217       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    196    196       P -> Q (IN REF. 2).
FT   CONFLICT    415    415       A -> R (IN REF. 2).
SQ   SEQUENCE   630 AA;  70147 MW;  F37EF1EC1764FB30 CRC64;

  Query Match              1.2%;  Score 9;  DB 1;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      99 LGNIWRFPY 107


RESULT 12
S6A4_RAT
ID   S6A4_RAT       STANDARD;      PRT;   630 AA.
AC   P31652; P23976;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium-dependent serotonin transporter (5HT transporter) (5HTT).
GN   SLC6A4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RX   MEDLINE=92111740; PubMed=1765155;
RA   Mayser W., Betz H., Schloss P.;
RT   "Isolation of cDNAs encoding a novel member of the neurotransmitter
RT   transporter gene family.";
RL   FEBS Lett. 295:203-206(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain stem;
RX   MEDLINE=92049754; PubMed=1944572;
RA   Blakely R.D., Berson H.E., Fremeau R.T. Jr., Caron M.G., Peek M.M.,
RA   Prince H.K., Bardley C.C.;
RT   "Cloning and expression of a functional serotonin transporter from
RT   rat brain.";
RL   Nature 354:66-70(1991).
RN   [3]
RP   REVISIONS.
RC   TISSUE=Brain stem;
RA   Blakely R.D.;
RL   Submitted (JUL-1992) to the EMBL/GenBank/DDBJ databases.
RN   [4]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92054541; PubMed=1948036;
RA   Hoffman B.J., Mezey E., Brownstein M.J.;
RT   "Cloning of a serotonin transporter affected by antidepressants.";
RL   Science 254:579-580(1991).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Fawn hooded;
RA   Gonzalez A.M., Smith A.P.L., Emery C.J., Higenbottam T.W.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: TERMINATES THE ACTION OF SEROTONINE BY ITS HIGH
CC       AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR .
CC       STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63995; CAA45401.1; -.
DR   EMBL; X63253; CAA44913.1; -.
DR   EMBL; M79450; AAA42186.1; -.
DR   EMBL; Y11024; CAA71909.1; -.
DR   PIR; S30604; S30604.
DR   PIR; S19585; S19585.
DR   InterPro; IPR002437; 5HT_transporter.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   Pfam; PF03491; 5HT_transporter; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport.
FT   DOMAIN        1     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    108       1 (POTENTIAL).
FT   TRANSMEM    116    135       2 (POTENTIAL).
FT   TRANSMEM    160    180       3 (POTENTIAL).
FT   DOMAIN      181    252       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    253    271       4 (POTENTIAL).
FT   TRANSMEM    280    297       5 (POTENTIAL).
FT   TRANSMEM    333    350       6 (POTENTIAL).
FT   TRANSMEM    362    383       7 (POTENTIAL).
FT   TRANSMEM    417    436       8 (POTENTIAL).
FT   TRANSMEM    464    482       9 (POTENTIAL).
FT   TRANSMEM    498    518       10 (POTENTIAL).
FT   TRANSMEM    539    558       11 (POTENTIAL).
FT   TRANSMEM    577    595       12 (POTENTIAL).
FT   DOMAIN      596    630       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    217    217       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    415    415       A -> G (IN REF. 4).
FT   CONFLICT    533    536       PGWF -> GMV (IN REF. 4).
FT   CONFLICT    621    630       PCGDIRMNAV -> RVGHPHECCVTHPGRGHLFPATSLSSE
FT                                KPTGLLL (IN REF. 4).
SQ   SEQUENCE   630 AA;  70171 MW;  44DA7C5888C403EE CRC64;

  Query Match             1.2%;  Score 9;  DB 1;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      99 LGNIWRFPY 107
```

```
RESULT 13
S6A7_HUMAN
ID   S6A7_HUMAN     STANDARD;      PRT;   636 AA.
AC   Q99884;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium-dependent proline transporter.
GN   SLC6A7 OR PROT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hippocampus;
RX   MEDLINE=95379780; PubMed=7651355;
RA   Shafqat S., Velaz-Faircloth M., Henzi V.A., Whitney K.D.,
RA   Yang-Feng T.L., Seldin M.F., Fremeau R.T. Jr.;
RT   "Human brain-specific L-proline transporter: molecular cloning,
RT   functional expression, and chromosomal localization of the gene in
RT   human and mouse genomes.";
RL   Mol. Pharmacol. 48:219-229(1995).
CC   -!- FUNCTION: TERMINATES THE ACTION OF PROLINE BY ITS HIGH AFFINITY
CC       SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S80071; AAB47007.2; -.
DR   Genew; HGNC:11054; SLC6A7.
DR   MIM; 606205; -.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Amino-acid transport; Polymorphism.
FT   DOMAIN        1     45       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     46     66       1 (POTENTIAL).
FT   TRANSMEM     74     93       2 (POTENTIAL).
FT   TRANSMEM    117    137       3 (POTENTIAL).
FT   DOMAIN      138    214       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    215    233       4 (POTENTIAL).
FT   TRANSMEM    242    259       5 (POTENTIAL).
FT   TRANSMEM    295    312       6 (POTENTIAL).
FT   TRANSMEM    324    345       7 (POTENTIAL).
FT   TRANSMEM    378    397       8 (POTENTIAL).
FT   TRANSMEM    425    443       9 (POTENTIAL).
FT   TRANSMEM    459    479       10 (POTENTIAL).
FT   TRANSMEM    500    519       11 (POTENTIAL).
FT   TRANSMEM    538    556       12 (POTENTIAL).
FT   DOMAIN      557    636       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    182    182       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     345    345       L -> V (IN DBSNP:1468564).
FT                                /FTId=VAR_011390.
SQ   SEQUENCE   636 AA;  70892 MW;  9A2D1071F59E409B CRC64;

  Query Match             1.2%;  Score 9;  DB 1;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLEL 118
           |||||||||
Db      87 GIPLFFLEL 95


RESULT 14
S6A7_RAT
ID   S6A7_RAT       STANDARD;      PRT;   637 AA.
AC   P28573;
DT   01-DEC-1992 (Rel. 24, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium-dependent proline transporter.
GN   SLC6A7.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=92265305; PubMed=1350201;
RA   Fremeau R.T. Jr., Caron M.G., Blakely R.D.;
RT   "Molecular cloning and expression of a high affinity L-proline
RT   transporter expressed in putative glutamatergic pathways of rat
RT   brain.";
RL   Neuron 8:915-926(1992).
CC   -!- FUNCTION: TERMINATES THE ACTION OF PROLINE BY ITS HIGH AFFINITY
CC       SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SUBPOPULATIONS OF PUTATIVE
CC       GLUTAMATERGIC PATHWAYS OF RAT BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC       FAMILY (SNF).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M88111; AAA41541.1; ALT_TERM.
DR   InterPro; IPR000175; Na/ntran_symport.
DR   Pfam; PF00209; SNF; 1.
DR   PRINTS; PR00176; NANEUSMPORT.
DR   ProDom; PD000448; Na/ntran_symport; 1.
DR   PROSITE; PS00610; NA_NEUROTRAN_SYMP_1; 1.
DR   PROSITE; PS00754; NA_NEUROTRAN_SYMP_2; 1.
DR   PROSITE; PS50267; NA_NEUROTRAN_SYMP_3; 1.
KW   Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW   Symport; Amino-acid transport.
FT   DOMAIN        1     45       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     46     66       1 (POTENTIAL).
FT   TRANSMEM     74     93       2 (POTENTIAL).
FT   TRANSMEM    117    137       3 (POTENTIAL).
FT   DOMAIN      138    214       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    215    233       4 (POTENTIAL).
FT   TRANSMEM    242    259       5 (POTENTIAL).
FT   TRANSMEM    295    312       6 (POTENTIAL).
FT   TRANSMEM    324    345       7 (POTENTIAL).
FT   TRANSMEM    378    397       8 (POTENTIAL).
FT   TRANSMEM    425    443       9 (POTENTIAL).
FT   TRANSMEM    459    479       10 (POTENTIAL).
FT   TRANSMEM    500    519       11 (POTENTIAL).
FT   TRANSMEM    538    556       12 (POTENTIAL).
FT   DOMAIN      557    637       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    182    182       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   637 AA;  71090 MW;  9627E6DD5BBC9408 CRC64;

  Query Match             1.2%;  Score 9;  DB 1;  Length 637;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLEL 118
           |||||||||
Db      87 GIPLFFLEL 95


RESULT 15
NU6M_DROYA
ID   NU6M_DROYA     STANDARD;      PRT;   174 AA.
AC   P07709;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
GN   ND6.
OS   Drosophila yakuba (Fruit fly).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2317.6 Ivory Coast;
RX   MEDLINE=86089137; PubMed=3001325;
RA   Clary D.O., Wolstenholme D.R.;
RT   "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide
RT   sequence, gene organization, and genetic code.";
RL   J. Mol. Evol. 22:252-271(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84221393; PubMed=6328435;
RA   Clary D.O., Wahleithner J.A., Wolstenholme D.R.;
RT   "Sequence and arrangement of the genes for cytochrome b, URF1, URF4L,
RT   URF4, URF5, URF6 and five tRNAs in Drosophila mitochondrial DNA.";
RL   Nucleic Acids Res. 12:3747-3762(1984).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X03240; CAA26995.1; -.
DR   PIR; B30020; B30020.
DR   FlyBase; FBgn0013187; Dyak\mt:ND6.
DR   InterPro; IPR001457; Oxidored_q3.
DR   Pfam; PF00499; oxidored_q3; 1.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ   SEQUENCE   174 AA;  20200 MW;  469DA42041574C57 CRC64;

  Query Match             1.1%;  Score 8;  DB 1;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     620 LLITLIVV 627
           ||||||||
Db     152 LLITLIVV 159


Search completed: April 21, 2003, 14:10:18
Job time : 17 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:08:39 ; Search time 25 Seconds
(without alignments)
2795.589 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 727
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	11.4	727	2	S27043	neurotransmitter t
2	83	11.4	727	2	I56506	Na+/Cl(-)-dependen
3	33	4.5	33	2	C46027	neurotransmitter t
4	22	3.0	728	2	I65413	sodium-dependent n
5	22	3.0	730	2	I52632	sodium-dependent n
6	11	1.5	499	2	S28306	hypothetical prote
7	11	1.5	505	2	G84079	sodium-dependent t
8	11	1.5	508	2	E64089	probable transport
9	11	1.5	511	2	E81964	probable sodium-de
10	11	1.5	511	2	H81019	sodium- and chlori
11	11	1.5	579	2	D88551	protein T23G5.5 [i
12	11	1.5	615	2	T43330	catecholamine tran
13	10	1.4	527	2	E82127	sodium-dependent t
14	10	1.4	635	2	T20910	hypothetical prote
15	10	1.4	706	2	T20052	hypothetical prote
16	9	1.2	453	2	H83790	sodium-dependent t
17	9	1.2	607	2	S19585	serotonin transpor
18	9	1.2	616	2	S50998	transporter protei
19	9	1.2	630	2	S30604	neurotransmitter t
20	9	1.2	637	2	JH0674	L-proline transpor
21	9	1.2	768	2	T22758	hypothetical prote
22	8	1.1	174	2	B30020	NADH2 dehydrogenas
23	8	1.1	264	2	D81253	probable peptide A
24	8	1.1	318	2	E82493	conserved hypothet
25	8	1.1	431	2	B75491	proton/sodium-glut
26	8	1.1	491	2	B69499	sodium- and chlori
27	8	1.1	492	2	F64464	sodium-dependent n
28	8	1.1	521	2	A99549	amino acid permeas
29	8	1.1	611	2	S19434	probable transport
30	8	1.1	690	2	S75067	H+/K+-exchanging A
31	8	1.1	743	2	H97302	ribonucleotide red
32	8	1.1	1017	2	T37201	hypothetical prote
33	7	1.0	27	2	G61002	transformation-sen
34	7	1.0	33	2	B46027	neurotransmitter t
35	7	1.0	33	2	A46027	probable noradrena
36	7	1.0	33	2	E46027	glycine transporte
37	7	1.0	66	2	D83599	conserved hypothet
38	7	1.0	67	2	C48716	glycine transporte
39	7	1.0	73	2	S38880	maltose-binding pr
40	7	1.0	102	2	AD1748	B. subtilis YfhL p
41	7	1.0	107	2	S50684	hypothetical prote
42	7	1.0	115	2	A82626	conserved hypothet
43	7	1.0	118	2	C83796	hypothetical prote
44	7	1.0	131	2	T05837	hypothetical prote
45	7	1.0	132	2	AC0780	probable membrane

ALIGNMENTS

RESULT 1
S27043
neurotransmitter transport protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S27043
R;Liu, Q.R.; Mandiyan, S.; Lopez-Corcuera, B.; Nelson, H.; Nelson, N.
FEBS Lett. 315, 114-118, 1993
A;Title: A rat brain cDNA encoding the neurotransmitter transporter with an unusual
A;Reference number: S27043; MUID:93114444; PMID:8093354
A;Accession: S27043
A;Molecule type: mRNA
A;Residues: 1-727 <LIU>
A;Cross-references: GB:S52051; NID:g262842; PIDN:AAB24776.1; PID:g262843
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: transmembrane protein

Query Match 11.4%; Score 83; DB 2; Length 727;
Best Local Similarity 100.0%; Pred. No. 4e-77;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111

Qy     112 PLFFLELAVGQRIRRGSIGVWHY 134
           |||||||||||||||||||||||
Db     112 PLFFLELAVGQRIRRGSIGVWHY 134
```

RESULT 2
I56506
Na+/Cl(-)-dependent neurotransmitter transporter, brain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C;Accession: I56506
R;el Mestikawy, S.; Giros, B.; Pohl, M.; Hamon, M.; Kingsmore, S.F.; Seldin, M.F.;
J. Neurochem. 62, 445-455, 1994
A;Title: Characterization of an atypical member of the Na+/Cl(-)-dependent transpo
A;Reference number: I56506; MUID:94125086; PMID:8294906
A;Accession: I56506
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-727 <RES>
A;Cross-references: GB:S68944; NID:g545077; PIDN:AAC60673.1; PID:g545078
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 11.4%; Score 83; DB 2; Length 727;
Best Local Similarity 100.0%; Pred. No. 4e-77;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111

Qy     112 PLFFLELAVGQRIRRGSIGVWHY 134
           |||||||||||||||||||||||
Db     112 PLFFLELAVGQRIRRGSIGVWHY 134


RESULT 3
C46027
neurotransmitter transporter - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Aug-1998
C;Accession: C46027
R;Liu, Q.R.; Mandiyan, S.; Nelson, H.; Nelson, N.
Proc. Natl. Acad. Sci. U.S.A. 89, 6639-6643, 1992
A;Title: A family of genes encoding neurotransmitter transporters.
A;Reference number: A46027; MUID:92335351; PMID:1631167
A;Accession: C46027
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-33 <LIU>
A;Note: sequence extracted from NCBI backbone (NCBIN:108827, NCBIP:108828)
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             4.5%;  Score 33;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-26;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      63 WNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 95
           |||||||||||||||||||||||||||||||||
Db       1 WNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGG 33


RESULT 4
I65413
sodium-dependent neurotransmitter transporter - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C;Accession: I65413
R;Uhl, G.R.; Kitayama, S.; Gregor, P.; Nanthakumar, E.; Persico, A.; Shimada, S.
Brain Res. Mol. Brain Res. 16, 353-359, 1992
A;Title: Neurotransmitter transporter family cDNAs in a rat midbrain library: 'orphan tr
A;Reference number: I52632; MUID:93180651; PMID:1363329
A;Accession: I65413
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-728 <RES>
A;Cross-references: GB:S56968; NID:g298323; PIDN:AAB25532.1; PID:g298324
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             3.0%;  Score 22;  DB 2;  Length 728;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     511 FVQRSGNYFVTMFDDYSATLPL 532


RESULT 5
I52632
sodium-dependent neurotransmitter transporter - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 21-Aug-1998
C;Accession: I52632
R;Uhl, G.R.; Kitayama, S.; Gregor, P.; Nanthakumar, E.; Persico, A.; Shimada, S.
Brain Res. Mol. Brain Res. 16, 353-359, 1992
A;Title: Neurotransmitter transporter family cDNAs in a rat midbrain library: 'orphan tr
A;Reference number: I52632; MUID:93180651; PMID:1363329
A;Accession: I52632
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-730 <RES>
A;Cross-references: GB:S56141; NID:g298325
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             3.0%;  Score 22;  DB 2;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     513 FVQRSGNYFVTMFDDYSATLPL 534


RESULT 6
S28306
hypothetical protein T23G5.5 - Caenorhabditis elegans (fragment)
C;Species: Caenorhabditis elegans
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Aug-1998
C;Accession: S28306
R;Berks, M.
submitted to the EMBL Data Library, December 1992
A;Reference number: S28296
A;Accession: S28306
A;Molecule type: DNA
A;Residues: 1-499 <BER>
A;Cross-references: EMBL:Z19158
C;Genetics:
A;Introns: 27/3; 134/2; 184/2; 215/3; 253/2; 332/3; 375/3; 409/1; 442/3
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             1.5%;  Score 11;  DB 2;  Length 499;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPYLC 90
           ||||||||||||
Db      41 LGNIWRFPYLC 51


RESULT 7
G84079
sodium-dependent transporter BH3439 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: G84079
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G84079
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB07158.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3439
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             1.5%;  Score 11;  DB 2;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||||
Db      20 VGLGNIWRFPY 30


RESULT 8
E64089
probable transport protein HI0736 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Aug-1999
```

```
C;Accession: E64089
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: E64089
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-508 <TIGR>
A;Cross-references: GB:U32757; GB:L42023; NID:g1573740; PIDN:AAC22395.1; PID:g1573742; T
C;Superfamily: gamma-aminobutyric acid transporter
C;Keywords: transmembrane protein

  Query Match              1.5%;  Score 11;  DB 2;  Length 508;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||| 
Db      29 VGLGNIWRFPY 39


RESULT 9
E81964
probable sodium-dependent inner membrane transport protein NMA0470 [imported] - Neisseri
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: E81964
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: E81964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <PAR>
A;Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83765.1; PID:g737921
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0470
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match              1.5%;  Score 11;  DB 2;  Length 511;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||| 
Db      27 VGLGNIWRFPY 37


RESULT 10
H81019
sodium- and chloride-dependent transporter NMB1975 [imported] - Neisseria meningitidis (
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: H81019
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: H81019
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <TET>
A;Cross-references: GB:AE002546; GB:AE002098; NID:g7227234; PIDN:AAF42303.1; PID:g722723
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1975
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match              1.5%;  Score 11;  DB 2;  Length 511;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||| 
Db      27 VGLGNIWRFPY 37


RESULT 11
D88551
protein T23G5.5 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: D88551
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 199
A;Accession: D88551
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-579 <STO>
A;Cross-references: GB:chr_III; PIDN:CAA79564.1; PID:g3879264; GSPDB:GN00021; CESP
A;Note: similar to sodium:neurotransmitter symporter
C;Genetics:
A;Gene: T23G5.5
A;Map position: 3
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match              1.5%;  Score 11;  DB 2;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPYLC 90
           ||||||||||| 
Db      41 LGNIWRFPYLC 51


RESULT 12
T43330
catecholamine transport protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T43330
R;Jayanthi, L.D.; Apparsundaram, S.; Malone, M.D.; Ward, E.; Miller, D.M.; Eppler,
Mol. Pharmacol. 54, 601-609, 1998
A;Title: The Caenorhabditis elegans gene T23G5.5 encodes an antidepressant- and co
A;Reference number: Z22431; MUID:98440631; PMID:9765501
A;Accession: T43330
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-615 <JAY>
A;Cross-references: EMBL:AF079899; NID:g3982784; PIDN:AAC83661.1; PID:g3982785
C;Genetics:
A;Note: T23G5.5
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match              1.5%;  Score 11;  DB 2;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPYLC 90
           ||||||||||| 
Db      60 LGNIWRFPYLC 70
```

RESULT 13
E82127
sodium-dependent transporter VC2012 [imported] - Vibrio cholerae (strain N16961 serogrou
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: E82127
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82127
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-527 <HEI>
A;Cross-references: GB:AE004276; GB:AE003852; NID:g9656555; PIDN:AAF95160.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2012
A;Map position: 1
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 1.4%; Score 10; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 GLGNIWRFPY 88
||||||||||
Db 63 GLGNIWRFPY 72

RESULT 14
T20910
hypothetical protein ZK1010.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T20910; T27646
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19345
A;Accession: T20910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-635 <WIL>
A;Cross-references: EMBL:Z81503; PIDN:CAB04115.1; GSPDB:GN00021; CESP:ZK1010.9
A;Experimental source: clone F14F7
R;Gardner, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20398
A;Accession: T27646
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-635 <WI2>
A;Cross-references: EMBL:Z82083; PIDN:CAB04975.1; GSPDB:GN00021; CESP:ZK1010.9
A;Experimental source: clone ZK1010
C;Genetics:
A;Gene: CESP:ZK1010.9
A;Map position: 3
A;Introns: 71/3; 184/2; 316/3; 403/1; 444/3; 547/3; 577/3
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 1.4%; Score 10; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 VGLGNIWRFP 87
||||||||||
Db 83 VGLGNIWRFP 92

RESULT 15
[redacted]

T20052
hypothetical protein C49C3.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T20052
R;Mortimore, B.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19218
A;Accession: T20052
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-706 <WIL>
A;Cross-references: EMBL:Z82264; PIDN:CAB05153.1; GSPDB:GN00022; CESP:C49C3.1
A;Experimental source: clone C49C3
C;Genetics:
A;Gene: CESP:C49C3.1
A;Map position: 4
A;Introns: 50/3; 80/1; 125/1; 206/1; 240/2; 287/3; 325/3; 365/3; 487/3; 521/1; 564/3
C;Superfamily: gamma-aminobutyric acid transporter

Query Match 1.4%; Score 10; DB 2; Length 706;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 VGLGNIWRFP 87
||||||||||
Db 62 VGLGNIWRFP 71

Search completed: April 21, 2003, 14:11:37
Job time : 28 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:11:10 ; Search time 43 Seconds
(without alignments)
1278.402 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 727
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 288829 seqs, 75613885 residues

Word size : 0

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	727	100.0	727	10	US-09-923-444A-2	Sequence 2, Appli
2	22	3.0	729	10	US-09-741-149-4	Sequence 4, Appli
3	22	3.0	730	9	US-10-156-239-5	Sequence 5, Appli
4	22	3.0	730	10	US-09-741-149-2	Sequence 2, Appli
5	22	3.0	730	10	US-09-795-693-5	Sequence 5, Appli
6	15	2.1	676	9	US-09-815-923-12	Sequence 12, Appl
7	14	1.9	599	10	US-09-861-846-4	Sequence 4, Appli
8	14	1.9	610	10	US-09-861-846-2	Sequence 2, Appli
9	11	1.5	579	9	US-09-738-626-4648	Sequence 4648, Ap
10	10	1.4	624	9	US-10-156-239-24	Sequence 24, Appl
11	10	1.4	624	10	US-09-795-693-24	Sequence 24, Appl
12	8	1.1	224	10	US-09-843-598-6	Sequence 6, Appli
13	8	1.1	270	10	US-09-925-302-677	Sequence 677, App
14	8	1.1	421	10	US-09-843-598-9	Sequence 9, Appli
15	8	1.1	587	9	US-09-815-923-4	Sequence 4, Appli
16	8	1.1	624	9	US-10-105-959-4	Sequence 4, Appli
17	8	1.1	671	10	US-09-843-598-5	Sequence 5, Appli
18	8	1.1	671	10	US-09-843-598-7	Sequence 7, Appli
19	7	1.0	12	9	US-09-817-144-2	Sequence 2, Appli
20	7	1.0	15	10	US-09-818-656A-21	Sequence 21, Appl
21	7	1.0	52	10	US-09-864-761-43363	Sequence 43363, A
22	7	1.0	67	10	US-09-864-761-47143	Sequence 47143, A
23	7	1.0	68	10	US-09-864-761-34614	Sequence 34614, A
24	7	1.0	73	9	US-09-796-692-1130	Sequence 1130, Ap
25	7	1.0	102	10	US-09-921-397-31	Sequence 31, Appl
26	7	1.0	112	10	US-09-919-473-10	Sequence 10, Appl
27	7	1.0	127	10	US-09-921-397-110	Sequence 110, App
28	7	1.0	136	9	US-09-764-868-710	Sequence 710, App
29	7	1.0	209	9	US-09-957-187-79	Sequence 79, Appl
30	7	1.0	260	10	US-09-921-397-109	Sequence 109, App
31	7	1.0	266	9	US-09-738-626-6457	Sequence 6457, Ap
32	7	1.0	270	9	US-09-796-753-126	Sequence 126, App
33	7	1.0	270	9	US-09-905-674-2	Sequence 2, Appli
34	7	1.0	270	9	US-09-957-187-16	Sequence 16, Appl
35	7	1.0	297	9	US-09-510-332-87	Sequence 87, Appl
36	7	1.0	300	10	US-09-921-397-107	Sequence 107, App
37	7	1.0	305	9	US-09-510-332-135	Sequence 135, App
38	7	1.0	314	9	US-09-510-332-44	Sequence 44, Appl
39	7	1.0	314	9	US-09-510-332-71	Sequence 71, Appl
40	7	1.0	315	10	US-09-881-752A-54	Sequence 54, Appl
41	7	1.0	319	10	US-09-886-055-95	Sequence 95, Appl
42	7	1.0	335	10	US-09-862-810-4	Sequence 4, Appli
43	7	1.0	348	10	US-09-731-872-247	Sequence 247, App
44	7	1.0	352	9	US-10-102-806-543	Sequence 543, App
45	7	1.0	361	9	US-09-989-442-126	Sequence 126, App

ALIGNMENTS

```
RESULT 1
US-09-923-444A-2
; Sequence 2, Application US/09923444A
; Patent No. US20020015980A1
;  GENERAL INFORMATION:
;       APPLICANT: LI, YI
;                  FLEISCHMANN, ROBERT
;       TITLE OF INVENTION: NEUROTRANSMITTER TRANSPORTER
;       NUMBER OF SEQUENCES: 6
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: MD
;            COUNTRY: US
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/923,444A
;            FILING DATE: 08-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/062,815
;            FILING DATE: 199-12-20
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Michele M. Wales
;            REGISTRATION NUMBER: 43,975
;            REFERENCE/DOCKET NUMBER: PF116
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (301) 309-8504
;            TELEFAX: (301) 309-8439
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 727 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
```

US-09-923-444A-2

Query Match 100.0%; Score 727; DB 10; Length 727;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600

Qy     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660

Qy     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720

Qy     721 STPESEL 727
          |||||||
Db     721 STPESEL 727
```

RESULT 2
US-09-741-149-4
; Sequence 4, Application US/09741149
; Patent No. US20020031800A1
; GENERAL INFORMATION:
; APPLICANT: LI, Zhenya et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000780
; CURRENT APPLICATION NUMBER: US/09/741,149
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-741-149-4

Query Match 3.0%; Score 22; DB 10; Length 729;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533
```

RESULT 3
US-10-156-239-5
; Sequence 5, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human T
; TITLE OF INVENTION: ATPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule,
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefo
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-156-239-5

Query Match 3.0%; Score 22; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533
```

RESULT 4
US-09-741-149-2
; Sequence 2, Application US/09741149
; Patent No. US20020031800A1
; GENERAL INFORMATION:
; APPLICANT: LI, Zhenya et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000780
; CURRENT APPLICATION NUMBER: US/09/741,149
; CURRENT FILING DATE: 2000-12-21

```
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;    LENGTH: 730
;    TYPE: PRT
;    ORGANISM: Human
US-09-741-149-2

  Query Match              3.0%;  Score 22;  DB 10;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533


RESULT 5
US-09-795-693-5
; Sequence 5, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION:  20685, 579, 17114, 23821, 33894, and
;  TITLE OF INVENTION:  32613, No. US20020068710A1el Human Transporters
;  FILE REFERENCE:  35800/209292
;  CURRENT APPLICATION NUMBER: US/09/795,693
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  60/185,906
;  PRIOR FILING DATE:  2000-02-29
;  NUMBER OF SEQ ID NOS:  42
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;    LENGTH: 730
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-795-693-5

  Query Match              3.0%;  Score 22;  DB 10;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533


RESULT 6
US-09-815-923-12
; Sequence 12, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
;  APPLICANT: Gill, Sarjeet S.
;  APPLICANT:  Ross, Linda S.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1el
;  TITLE OF INVENTION:  Target Sites for Insecticides
;  FILE REFERENCE: 023070-093800US
;  CURRENT APPLICATION NUMBER: US/09/815,923
;  CURRENT FILING DATE:  2001-03-23
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;    LENGTH: 676
;    TYPE: PRT
;    ORGANISM: Manduca sexta
;    FEATURE:
;    OTHER INFORMATION: orphan transporter
US-09-815-923-12

  Query Match              2.1%;  Score 15;  DB 9;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      77 SVGLGNIWRFPYLCQ 91
           |||||||||||||||
Db      69 SVGLGNIWRFPYLCQ 83


RESULT 7
US-09-861-846-4
; Sequence 4, Application US/09861846
; Patent No. US20020110852A1
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al.
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001065
;  CURRENT APPLICATION NUMBER: US/09/861,846
;  CURRENT FILING DATE:  2001-05-22
;  PRIOR APPLICATION NUMBER: 09/752,821
;  PRIOR FILING DATE: 2001-01-03
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;    LENGTH: 599
;    TYPE: PRT
;    ORGANISM: Mus musculus
US-09-861-846-4

  Query Match              1.9%;  Score 14;  DB 10;  Length 599;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLCQ 91
           ||||||||||||||
Db      35 VGLGNIWRFPYLCQ 48


RESULT 8
US-09-861-846-2
; Sequence 2, Application US/09861846
; Patent No. US20020110852A1
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al.
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001065
;  CURRENT APPLICATION NUMBER: US/09/861,846
;  CURRENT FILING DATE:  2001-05-22
;  PRIOR APPLICATION NUMBER: 09/752,821
;  PRIOR FILING DATE: 2001-01-03
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;    LENGTH: 610
;    TYPE: PRT
;    ORGANISM: Human
US-09-861-846-2

  Query Match              1.9%;  Score 14;  DB 10;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPYLCQ 91
           ||||||||||||||
Db      36 VGLGNIWRFPYLCQ 49


RESULT 9
US-09-738-626-4648
; Sequence 4648, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4648
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-4648

  Query Match             1.5%;  Score 11;  DB 9;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 0.047;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
           ||||||||||
Db      52 VGLGNIWRFPY 62


RESULT 10
US-10-156-239-24
; Sequence 24, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Trans
;  TITLE OF INVENTION:  ATPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A Hu
;  TITLE OF INVENTION:  Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
;  FILE REFERENCE: 35800/247645
;  CURRENT APPLICATION NUMBER: US/10/156,239
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: 09/795,693
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/185,906
;  PRIOR FILING DATE: 2000-02-29
;  PRIOR APPLICATION NUMBER: 09/809,557
;  PRIOR FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/192,018
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 09/808,568
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/191,790
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 09/808,767
;  PRIOR FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/191,781
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 624
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-24

  Query Match             1.4%;  Score 10;  DB 9;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 0.46;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLELA 119
           ||||||||||
Db      51 GIPLFFLELA 60


RESULT 11
US-09-795-693-24
; Sequence 24, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION:  20685, 579, 17114, 23821, 33894, and
;  TITLE OF INVENTION:  32613, No. US20020068710A1el Human Transporters
;  FILE REFERENCE:  35800/209292
;  CURRENT APPLICATION NUMBER: US/09/795,693
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  60/185,906
;  PRIOR FILING DATE:  2000-02-29
;  NUMBER OF SEQ ID NOS:  42
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 624
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-24

  Query Match             1.4%;  Score 10;  DB 10;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 0.46;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLELA 119
           ||||||||||
Db      51 GIPLFFLELA 60


RESULT 12
US-09-843-598-6
; Sequence 6, Application US/09843598
; Patent No. US20020010944A1
; GENERAL INFORMATION:
;  APPLICANT: Horvitz, H. Robert
;  APPLICANT:  Ranganathan, Rajesh
;  TITLE OF INVENTION: CeSERT GENES, PROTEINS, AND MODULATORY
;  TITLE OF INVENTION:  COMPOUNDS
;  FILE REFERENCE: 01997/525002
;  CURRENT APPLICATION NUMBER: US/09/843,598
;  CURRENT FILING DATE:  2001-04-26
;  PRIOR APPLICATION NUMBER: US 60/200,549
;  PRIOR FILING DATE: 2000-04-26
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-09-843-598-6

  Query Match             1.1%;  Score 8;  DB 10;  Length 224;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFP 87
           ||||||||
Db     120 LGNIWRFP 127
```

```
RESULT 13
US-09-925-302-677
; Sequence 677, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 677
;   LENGTH: 270
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-677

  Query Match             1.1%;  Score 8;  DB 10;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    476 GTMAGITT 483
          ||||||||
Db     21 GTMAGITT 28


RESULT 14
US-09-843-598-9
; Sequence 9, Application US/09843598
; Patent No. US20020010944A1
; GENERAL INFORMATION:
;  APPLICANT: Horvitz, H. Robert
;  APPLICANT:  Ranganathan, Rajesh
;  TITLE OF INVENTION: CeSERT GENES, PROTEINS, AND MODULATORY
;  TITLE OF INVENTION:  COMPOUNDS
;  FILE REFERENCE: 01997/525002
;  CURRENT APPLICATION NUMBER: US/09/843,598
;  CURRENT FILING DATE:  2001-04-26
;  PRIOR APPLICATION NUMBER: US 60/200,549
;  PRIOR FILING DATE: 2000-04-26
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 421
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-09-843-598-9

  Query Match             1.1%;  Score 8;  DB 10;  Length 421;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     80 LGNIWRFP 87
          ||||||||
Db    120 LGNIWRFP 127


RESULT 15
US-09-815-923-4
; Sequence 4, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
;  APPLICANT: Gill, Sarjeet S.
;  APPLICANT:  Ross, Linda S.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1el
;  TITLE OF INVENTION:  Target Sites for Insecticides
;  FILE REFERENCE: 02307O-093800US
;  CURRENT APPLICATION NUMBER: US/09/815,923
;  CURRENT FILING DATE:  2001-03-23
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Manduca sexta
;   FEATURE:
;   OTHER INFORMATION: serotonin transporter
US-09-815-923-4

  Query Match             1.1%;  Score 8;  DB 9;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    112 PLFFLELA 119
          ||||||||
Db     90 PLFFLELA 97


Search completed: April 21, 2003, 14:19:57
Job time : 45 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:09:29 ; Search time 18 Seconds
(without alignments)
1188.359 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 727
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	540	74.3	727	1	US-08-424-424B-2	Sequence 2, Appli
2	540	74.3	727	5	PCT-US94-05363A-2	Sequence 2, Appli
3	11	1.5	247	4	US-09-134-001C-2877	Sequence 2877, Ap
4	9	1.2	607	1	US-07-959-943-7	Sequence 7, Appli
5	9	1.2	607	1	US-07-879-617A-12	Sequence 12, Appl
6	9	1.2	607	1	US-08-753-985-12	Sequence 12, Appl
7	9	1.2	616	1	US-08-149-100-2	Sequence 2, Appli
8	9	1.2	630	1	US-07-959-943-9	Sequence 9, Appli
9	9	1.2	635	1	US-07-879-617A-9	Sequence 9, Appli
10	9	1.2	635	1	US-08-753-985-9	Sequence 9, Appli
11	9	1.2	653	1	US-07-782-298-2	Sequence 2, Appli
12	9	1.2	667	1	US-07-879-617A-8	Sequence 8, Appli
13	9	1.2	667	1	US-08-753-985-8	Sequence 8, Appli
14	8	1.1	1278	4	US-09-462-136-2	Sequence 2, Appli
15	7	1.0	8	4	US-08-444-818-671	Sequence 671, App
16	7	1.0	11	1	US-08-707-873-1	Sequence 1, Appli
17	7	1.0	11	2	US-08-707-874-1	Sequence 1, Appli
18	7	1.0	11	2	US-08-315-067-1	Sequence 1, Appli
19	7	1.0	11	3	US-09-193-797-1	Sequence 1, Appli
20	7	1.0	11	5	PCT-US95-01770-1	Sequence 1, Appli
21	7	1.0	12	4	US-09-146-755-2	Sequence 2, Appli
22	7	1.0	20	2	US-07-894-063A-28	Sequence 28, Appl
23	7	1.0	86	4	US-09-134-001C-3928	Sequence 3928, Ap
24	7	1.0	152	4	US-08-444-818-62	Sequence 62, Appl
25	7	1.0	187	1	US-07-853-985A-2	Sequence 2, Appli
26	7	1.0	187	1	US-07-681-703B-2	Sequence 2, Appli
27	7	1.0	187	1	US-08-184-236-2	Sequence 2, Appli
28	7	1.0	187	2	US-08-407-410B-2	Sequence 2, Appli
29	7	1.0	187	2	US-08-485-500-2	Sequence 2, Appli
30	7	1.0	187	5	PCT-US91-02370-2	Sequence 2, Appli
31	7	1.0	187	5	PCT-US94-04174-2	Sequence 2, Appli
32	7	1.0	244	4	US-09-134-001C-4735	Sequence 4735, Ap
33	7	1.0	262	4	US-09-134-001C-5418	Sequence 5418, Ap
34	7	1.0	273	4	US-09-134-001C-3016	Sequence 3016, Ap
35	7	1.0	305	1	US-08-266-451B-26	Sequence 26, Appl
36	7	1.0	305	2	US-08-748-725-26	Sequence 26, Appl
37	7	1.0	312	1	US-08-240-783B-6	Sequence 6, Appli
38	7	1.0	312	3	US-09-084-813-6	Sequence 6, Appli
39	7	1.0	312	5	PCT-US92-09662-6	Sequence 6, Appli
40	7	1.0	361	3	US-09-028-934-30	Sequence 30, Appl
41	7	1.0	387	4	US-08-867-611-12	Sequence 12, Appl
42	7	1.0	387	5	PCT-US92-06965A-17	Sequence 17, Appl
43	7	1.0	410	4	US-08-725-758A-4	Sequence 4, Appli
44	7	1.0	420	4	US-09-134-001C-3805	Sequence 3805, Ap
45	7	1.0	426	4	US-08-725-758A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-424-424B-2
; Sequence 2, Application US/08424424B
; Patent No. 5759854
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Neurotransmitter Transporter
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,424B
; FILING DATE: APRIL 21, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05363
; FILING DATE: MAY 25, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 727 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-424-424B-2

Query Match 74.3%; Score 540; DB 1; Length 727;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 188 SVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQ 247

```
Db     188 SVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQ 247

Qy     248 SSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFF 307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     248 SSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFF 307

Qy     308 GLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCV 367
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     308 GLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCV 367

Qy     368 VENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLL 427
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     368 VENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLL 427

Qy     428 EDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID 487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     428 EDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID 487

Qy     488 TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIY 547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     488 TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIY 547

Qy     548 GPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAA 607
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     548 GPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAA 607

Qy     608 ERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDISNL 667
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     608 ERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDISNL 667

Qy     668 EENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLASTPESEL 727
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     668 EENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLASTPESEL 727


RESULT 2
PCT-US94-05363A-2
; Sequence 2, Application PC/TUS9405363A
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Neurotransmitter Transporter
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05363A
;      FILING DATE:  SUBMITTED HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-118
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  727 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
PCT-US94-05363A-2

  Query Match             74.3%;  Score 540;  DB 5;  Length 727;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     188 SVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQ 247
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     188 SVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGMAVGKGIQ 247

Qy     248 SSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFF 307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     248 SSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWREVATQVFF 307

Qy     308 GLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCV 367
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     308 GLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKANIMNEKCV 367

Qy     368 VENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLL 427
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     368 VENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSALGLDPCLL 427

Qy     428 EDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID 487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     428 EDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAGITTPIID 487

Qy     488 TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIY 547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     488 TSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIY 547

Qy     548 GPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAA 607
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     548 GPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSAWIKEEAA 607

Qy     608 ERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDISNL 667
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     608 ERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARMMKDISNL 667

Qy     668 EENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLASTPESEL 727
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     668 EENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLASTPESEL 727


RESULT 3
US-09-134-001C-2877
; Sequence 2877, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2877
;   LENGTH: 247
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2877

  Query Match             1.5%;  Score 11;  DB 4;  Length 247;
  Best Local Similarity   100.0%;  Pred. No. 0.0098;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      78 VGLGNIWRFPY 88
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/753,985
;      FILING DATE:  03-DEC-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/879617
;      FILING DATE:  01-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU109
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  607 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus
;      TISSUE TYPE:  Brain - serotonin transporter
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  95..96
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  102..103
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  109..110
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  116..117
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
US-08-753-985-12

  Query Match             1.2%;  Score 9;  DB 1;  Length 607;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      76 LGNIWRFPY 84


RESULT 7
US-08-149-100-2
; Sequence 2, Application US/08149100
; Patent No. 5559021
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kelli E.
;    APPLICANT:  Weinshank, Richard L.
;    TITLE OF INVENTION:  DNA ENCODING A NOVEL MAMMALIAN
;    TITLE OF INVENTION:  TRANSPORTER HOMOLOGOUS TO NEUROTRANSMITTER TRANSPORTERS
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/149,100
;      FILING DATE:  08-NOV-1993
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  44744/JPW/TEP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 977-9550
;      TELEFAX:  (212) 6640525
;      TELEX:  422523 COOP UI
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  616 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-149-100-2

  Query Match             1.2%;  Score 9;  DB 1;  Length 616;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     435 VQGTGLAFI 443
           |||||||||
Db     392 VQGTGLAFI 400


RESULT 8
US-07-959-943-9
; Sequence 9, Application US/07959943
; Patent No. 5418162
;  GENERAL INFORMATION:
;    APPLICANT:  Blakely, Randy D.
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    TITLE OF INVENTION:  Serotonin Transporter cDNA
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenneth D. Sibley; Bell, Seltzer, Park
;      ADDRESSEE:  and
;      ADDRESSEE:  Gibson
;      STREET:  Post Office Drawer 31107
;      CITY:  Raleigh
;      STATE:  No. 5418162th Carolina
;      COUNTRY:  U.S.A.
;      ZIP:  27622
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/959,943
;      FILING DATE:  19921014
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sibley, Kenneth D.
;      REGISTRATION NUMBER:  31,665
;      REFERENCE/DOCKET NUMBER:  5405.38a
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-881-3140
;      TELEFAX:  919-881-3175
```

```
              ||||||||||
Db      55 VGLGNIWRFPY 65


RESULT 4
US-07-959-943-7
; Sequence 7, Application US/07959943
; Patent No. 5418162
;  GENERAL INFORMATION:
;    APPLICANT:  Blakely, Randy D.
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    TITLE OF INVENTION:  Serotonin Transporter cDNA
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenneth D. Sibley; Bell, Seltzer, Park
;      ADDRESSEE:  and
;      ADDRESSEE:  Gibson
;      STREET:  Post Office Drawer 31107
;      CITY:  Raleigh
;      STATE:  No. 5418162th Carolina
;      COUNTRY:  U.S.A.
;      ZIP:  27622
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/959,943
;      FILING DATE:  19921014
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sibley, Kenneth D.
;      REGISTRATION NUMBER:  31,665
;      REFERENCE/DOCKET NUMBER:  5405.38a
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-881-3140
;      TELEFAX:  919-881-3175
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  607 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-959-943-7

  Query Match             1.2%;  Score 9;  DB 1;  Length 607;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      76 LGNIWRFPY 84


RESULT 5
US-07-879-617A-12
; Sequence 12, Application US/07879617A
; Patent No. 5580775
;  GENERAL INFORMATION:
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    APPLICANT:  Blakely, Randy D.
;    TITLE OF INVENTION:  A High Affinity L-Proline Transporter
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  U.S.
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/879,617A
;      FILING DATE:  19920501
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU109
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  607 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus
;      TISSUE TYPE:  Brain - serotonin transporter
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  95..96
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  102..103
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  109..110
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  116..117
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
US-07-879-617A-12

  Query Match             1.2%;  Score 9;  DB 1;  Length 607;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      76 LGNIWRFPY 84


RESULT 6
US-08-753-985-12
; Sequence 12, Application US/08753985
; Patent No. 5759788
;  GENERAL INFORMATION:
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    APPLICANT:  Blakely, Randy D.
;    TITLE OF INVENTION:  A High Affinity L-Proline Transporter
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  U.S.
;      ZIP:  30309
```

```
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  630 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-959-943-9

  Query Match             1.2%;  Score 9;  DB 1;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      80 LGNIWRFPY 88
           |||||||||
Db      99 LGNIWRFPY 107


RESULT 9
US-07-879-617A-9
; Sequence 9, Application US/07879617A
; Patent No. 5580775
;  GENERAL INFORMATION:
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    APPLICANT:  Blakely, Randy D.
;    TITLE OF INVENTION:  A High Affinity L-Proline Transporter
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  U.S.
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/879,617A
;      FILING DATE:  19920501
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU109
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  635 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus
;      TISSUE TYPE:  Brain - Proline Transporter
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  46..65
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  72..97
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  117..137
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  217..236
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  243..264
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  291..311
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  322..345
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  373..397
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  424..443
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  456..479
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  500..519
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  536..559
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  76..77
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  83..84
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  90..91
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  97..98
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
US-07-879-617A-9

  Query Match             1.2%;  Score 9;  DB 1;  Length 635;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLEL 118
```

```
          |||||||||
Db      87 GIPLFFLEL 95


RESULT 10
US-08-753-985-9
; Sequence 9, Application US/08753985
; Patent No. 5759788
;  GENERAL INFORMATION:
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    APPLICANT:  Blakely, Randy D.
;    TITLE OF INVENTION:  A High Affinity L-Proline Transporter
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  U.S.
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/753,985
;      FILING DATE:  03-DEC-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/879617
;      FILING DATE:  01-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU109
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  635 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus
;      TISSUE TYPE:  Brain - Proline Transporter
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  46..65
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  72..97
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  117..137
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  217..236
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  243..264
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  291..311
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  322..345
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  373..397
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  424..443
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  456..479
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  500..519
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  536..559
;      OTHER INFORMATION:  /note= "Proposed transmembrane
;      OTHER INFORMATION:  domain."
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  76..77
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  83..84
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  90..91
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Active-site
;      LOCATION:  97..98
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
US-08-753-985-9

  Query Match             1.2%;  Score 9;  DB 1;  Length 635;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLEL 118
           |||||||||
Db      87 GIPLFFLEL 95


RESULT 11
US-07-782-298-2
; Sequence 2, Application US/07782298
; Patent No. 5552308
;  GENERAL INFORMATION:
;    APPLICANT:  Hoffman, Beth.J.
```

; APPLICANT: Mezey, Eva
; APPLICANT: Brownstein, Michael J.
; TITLE OF INVENTION: cDNA Clone of a Rat Serotonin
; TITLE OF INVENTION: Transporter and Protein Encoded Thereby
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/782,298
; FILING DATE: October 24, 1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gerald M. Murphy, Jr.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-782-298-2

Query Match 1.2%; Score 9; DB 1; Length 653;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 LGNIWRFPY 88
|||||||||
Db 99 LGNIWRFPY 107

RESULT 12
US-07-879-617A-8
; Sequence 8, Application US/07879617A
; Patent No. 5580775
; GENERAL INFORMATION:
; APPLICANT: Fremeau Jr., Robert T.
; APPLICANT: Caron, Marc G.
; APPLICANT: Blakely, Randy D.
; TITLE OF INVENTION: A High Affinity L-Proline Transporter
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/879,617A
; FILING DATE: 19920501
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 667 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rattus
; TISSUE TYPE: Brain
; IMMEDIATE SOURCE:
; LIBRARY: rat forebrain cDNA library
; CLONE: rTB2-2-20
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 76..95
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 103..127
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 147..167
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 247..266
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 272..294
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 321..341
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 353..375
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 404..427
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 454..473
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 487..509
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 530..549
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 567..589
; OTHER INFORMATION: /note= "Membrane-spanning domain"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 44..45
; OTHER INFORMATION: /note= "protein kinase C
; OTHER INFORMATION: phosphorylation site"

```
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  71..72
;      OTHER INFORMATION:  /note= "cAMP-dependent protein
;      OTHER INFORMATION:  kinase phosphorylation site"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  269..270
;      OTHER INFORMATION:  /note= "protein kinase C
;      OTHER INFORMATION:  phosphorylation site"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  630..631
;      OTHER INFORMATION:  /note= "protein kinase C
;      OTHER INFORMATION:  phosphorylation site"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  106..127
;      OTHER INFORMATION:  /note= "Leucine zipper motif"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  212..213
;      OTHER INFORMATION:  /note= "N-linked glycosylation
;      OTHER INFORMATION:  site"
US-07-879-617A-8

  Query Match             1.2%;  Score 9;  DB 1;  Length 667;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLEL 118
           |||||||||
Db     117 GIPLFFLEL 125


RESULT 13
US-08-753-985-8
; Sequence 8, Application US/08753985
; Patent No. 5759788
;  GENERAL INFORMATION:
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    APPLICANT:  Blakely, Randy D.
;    TITLE OF INVENTION:  A High Affinity L-Proline Transporter
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  U.S.
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/753,985
;      FILING DATE:  03-DEC-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/879617
;      FILING DATE:  01-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU109
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  667 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Rattus
;      TISSUE TYPE:  Brain
;    IMMEDIATE SOURCE:
;      LIBRARY:  rat forebrain cDNA library
;      CLONE:  rTB2-2-20
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  76..95
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  103..127
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  147..167
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  247..266
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  272..294
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  321..341
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  353..375
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  404..427
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  454..473
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  487..509
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  530..549
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  567..589
;      OTHER INFORMATION:  /note= "Membrane-spanning domain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  44..45
;      OTHER INFORMATION:  /note= "protein kinase C
;      OTHER INFORMATION:  phosphorylation site"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  71..72
;      OTHER INFORMATION:  /note= "cAMP-dependent protein
;      OTHER INFORMATION:  kinase phosphorylation site"
;    FEATURE:
```

```
;       NAME/KEY:  Region
;       LOCATION:  269..270
;       OTHER INFORMATION:  /note= "protein kinase C
;       OTHER INFORMATION:  phosphorylation site"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  630..631
;       OTHER INFORMATION:  /note= "protein kinase C
;       OTHER INFORMATION:  phosphorylation site"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  106..127
;       OTHER INFORMATION:  /note= "Leucine zipper motif"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  212..213
;       OTHER INFORMATION:  /note= "N-linked glycosylation
;       OTHER INFORMATION:  site"
US-08-753-985-8

  Query Match             1.2%;  Score 9;  DB 1;  Length 667;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     110 GIPLFFLEL 118
           |||||||||
Db     117 GIPLFFLEL 125


RESULT 14
US-09-462-136-2
; Sequence 2, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
;  APPLICANT:  Carstea, et al.
;  TITLE OF INVENTION:  Genes for Niemann-Pick Type C Disease
;  FILE REFERENCE:  4239-53894
;  CURRENT APPLICATION NUMBER: US/09/462,136
;  CURRENT FILING DATE:  2000-06-01
;  PRIOR APPLICATION NUMBER:  PCT/US98/13862
;  PRIOR FILING DATE:  1998-07-02
;  PRIOR APPLICATION NUMBER:  US 60/051,682
;  PRIOR FILING DATE:  1997-07-03
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1278
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-462-136-2

  Query Match             1.1%;  Score 8;  DB 4;  Length 1278;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     530 LPLTLIVI 537
           ||||||||
Db     680 LPLTLIVI 687


RESULT 15
US-08-444-818-671
; Sequence 671, Application US/08444818
; Patent No. 6150087
;  GENERAL INFORMATION:
;    APPLICANT:  Chien, David Y.
;    APPLICANT:  Rutter, William J.
;    TITLE OF INVENTION:  NANBV Diagnostics and Vaccines
;    NUMBER OF SEQUENCES:  777
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608-2916
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,818
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/403,590
;      FILING DATE:  14-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Harbin, Alisa A.
;      REGISTRATION NUMBER:  33,895
;      REFERENCE/DOCKET NUMBER:  0110.002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (508)359-3876
;      TELEFAX:  (508)359-3885
;  INFORMATION FOR SEQ ID NO:  671:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-444-818-671

  Query Match             1.0%;  Score 7;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     444 AFTEAMT 450
           |||||||
Db       1 AFTEAMT 7


Search completed: April 21, 2003, 14:12:01
Job time : 21 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:05:49 ; Search time 41 Seconds
(without alignments)
2362.761 Million cell updates/sec

Title: US-09-923-444A-2
Perfect score: 727
Sequence: 1 MPKNSKVTQREHSSEHVTES.....NPNGPYGRGYLLASTPESEL 727

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	626	86.1	727	17	AAR88390	Human neurotransmi
2	182	25.0	727	23	AAE21800	Human HIPHUM 00000
3	56	7.7	266	23	ABB06153	Human NS protein s
4	52	7.2	122	22	AAO05876	Human polypeptide
5	44	6.1	143	22	ABB11188	Human neurotransmi
6	37	5.1	127	22	AAM24377	Rat EST encoded pr
7	26	3.6	83	22	AAM24378	Rat EST encoded pr
8	22	3.0	185	22	ABB11163	Human orphan trans
9	22	3.0	188	23	ABB57432	Human secreted pro
10	22	3.0	188	23	ABB57433	Human secreted pro
11	22	3.0	259	23	AAO21214	Sodium-dependent n
12	22	3.0	392	22	AAG64743	Human sodium neuro
13	22	3.0	729	23	AAO21215	Protein sequence i
14	22	3.0	730	22	AAG67159	Amino acid sequenc
15	22	3.0	730	22	AAY72908	Human NTT7 protein
16	22	3.0	730	23	ABB77168	Human GABA transpo
17	22	3.0	730	23	AAE14404	Human neurotransmi
18	19	2.6	249	22	AAB92689	Human protein sequ
19	18	2.5	136	22	ABB11223	Human orphan trans
20	16	2.2	289	22	AAB94304	Human protein sequ
21	15	2.1	675	22	AAE05100	Drosophila melanog
22	15	2.1	744	22	ABB61363	Drosophila melanog
23	14	1.9	259	22	ABG23381	Novel human diagno
24	14	1.9	259	22	ABG23431	Novel human diagno
25	12	1.7	12	20	AAY32821	Sodium-, chloride-
26	11	1.5	247	23	ABP38032	Staphylococcus epi
27	11	1.5	318	22	AAU57896	Propionibacterium
28	11	1.5	579	22	AAG90894	C glutamicum prote
29	11	1.5	596	20	AAY09518	C. elegans dopamin
30	11	1.5	615	20	AAY09517	C. elegans dopamin
31	9	1.2	592	23	AAE14408	Human neurotransmi
32	9	1.2	607	14	AAR34662	Rat 5HT transporte
33	9	1.2	607	16	AAR76072	Rat 5HT transporte
34	9	1.2	616	17	AAW07635	Rat brain derived
35	9	1.2	616	22	AAG64193	Human nerve mass-t
36	9	1.2	630	14	AAR34663	Rat 5HT transporte
37	9	1.2	630	16	AAR76073	Rat 5HT transporte
38	9	1.2	637	18	AAW11704	High affinity Na+-
39	9	1.2	647	20	AAW73376	Human HPDDV78 prot
40	9	1.2	653	13	AAR25642	5HTT serotonin tra
41	9	1.2	667	19	AAW57224	Rat proline transp
42	8	1.1	101	23	ABP33274	Human transport pr
43	8	1.1	106	22	ABG18975	Novel human diagno
44	8	1.1	116	23	ABB89187	Human polypeptide
45	8	1.1	130	20	AAY12151	Human 5' EST secre

ALIGNMENTS

RESULT 1
AAR88390
ID AAR88390 standard; Protein; 727 AA.
XX
AC AAR88390;
XX
DT 15-OCT-1996 (first entry)
XX
DE Human neurotransmitter transporter protein.
XX
KW Neurotransmitter transporter protein; pain therapy; stroke therapy;
KW amyotrophic lateral sclerosis.
XX
OS Homo sapiens.
XX
PN WO9531539-A1.
XX
PD 23-NOV-1995.
XX
PF 16-MAY-1994; 94WO-US05363.
XX
PR 16-MAY-1994; 94WO-US05363.
PR 26-MAY-1994; 94ZA-0003696.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Fleischmann RD, Li Y;
XX
DR WPI; 1996-010925/01.
DR N-PSDB; AAT09866.
XX
PT DNA encoding neuro:transmitter transporter protein and related
PT (ant)agonists - useful to treat e.g., amyotrophic lateral sclerosis,

```
PT   pain and stroke.
XX
PS   Claim 1; Fig.1; 53pp; English.
XX
CC   This neurotransmitter transporter protein may be expressed
CC   recombinantly in a host microorganism using a vector plasmid.
CC   NTT antagonists may be used to treat amyotrophic lateral sclerosis,
CC   pain and strokes.  NTT antagonists/inhibitors may be used to treat
CC   neurological and psychiatric disorders, e.g. depression, anxiety
CC   and epilepsy.
XX
SQ   Sequence   727 AA;

  Query Match             86.1%;  Score 626;  DB 17;  Length 727;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  726;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELDAEDR 60

Qy      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGIPLFFLELAV 120

Qy     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKSFQYPLPWSE 180

Qy     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCLLVVWSIGGM 240

Qy     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AVGKGIQSSGKVMYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKLVKMLDPQVWRE 300

Qy     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VATQVFFGLGLGFGGVIVFSSYNKQDNNCHFDGALVSFINFFTSVLATLVVFVVLGFKAN 360

Qy     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 IMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEMDNVIMTVKEDQFSAL 420

Qy     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GLDPCLLEDELDKSVQGTGLAFIAFTEAMTHFPTSPFWSVMFFLMLINLGLGSMIGTMAG 480

Qy     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ITTPIIDTSKVPKEMFTVGCCVFTFLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILEN 540

Qy     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 IAVAWIYGPKKFMQELTEMLGFRPYRFYFYMWKFVSPLCMAVLTTASIIQLGVTPPAYSA 600

Qy     601 WIKEEAAERYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 WIKEEAAEPYLYFPNWPMALLITLIVVATLPIPVVFVLRHFHLLSDGSNTLSVSYKKARM 660

Qy     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 MKDISNLEENDETRFILSKVPSEAPSPMPTHRSYLGPGSTSPLETSWNPNGPYGRGYLLA 720

Qy     721 STPESEL 727
           |||||||
Db     721 STPESEL 727


RESULT 2
AAE21800
ID   AAE21800 standard; Protein; 727 AA.
XX
AC   AAE21800;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Human HIPHUM 0000029 protein.
XX
KW   Human; neurotransmitter transporter polypeptide; HIPHUM 0000029 protein;
KW   prophylaxis; psychiatric disorder; bipolar disorder; psychotic disorder;
KW   unipolar depression; anxiety; schizophrenia; neurodegenerative disorder;
KW   neurological disorder; drug dependence; gene therapy; chromosome 1p21.
XX
OS   Homo sapiens.
XX
PN   GB2365432-A.
XX
PD   20-FEB-2002.
XX
PF   19-MAY-2000; 2000GB-0012199.
XX
PR   19-MAY-2000; 2000GB-0012199.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Sala CF,  Terstappen GC;
XX
DR   WPI; 2002-317841/36.
DR   N-PSDB; AAD34007.
XX
PT   An isolated neurotransmitter transporter polypeptide HIPHUM 0000029,
PT   whose regulation is useful in treating or preventing disease such as
PT   anxiety or depression -
XX
PS   Claim 1; Page 25-28; 39pp; English.
XX
CC   The invention relates to human neurotransmitter transporter polypeptides
CC   referred to as HIPHUM 0000029 and nucleic acid molecules encoding such
CC   polypeptides. Sequences of the invention their agonists and antagonists
CC   are useful for preparing a medicament for treatment or prophylaxis of
CC   disorders such as psychiatric disorders, bipolar disorders, unipolar
CC   depression, anxiety, schizophrenia, psychotic disorders, neurological/
CC   neurodegenerative disorders and drug dependence. Polynucleotides of the
CC   invention are useful as primers and probes. They are also used in gene
CC   therapy. The present sequence is human HIPHUM 0000029 protein. The
CC   HIPHUM 0000029 gene is located on chromosome 1p21.
XX
SQ   Sequence   727 AA;

  Query Match             25.0%;  Score 182;  DB 23;  Length 727;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-179;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52 EEELDAEDRPAWNSKLQYILAQIGFSVGLGNIWRFPYLCQKNGGGAYLVPYLVLLIIIGI 111

Qy     112 PLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKS 171
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 PLFFLELAVGQRIRRGSIGVWHYICPRLGGIGFSSCIVCLFVGLYYNVIIGWSIFYFFKS 171

Qy     172 FQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCL 231
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     172 FQYPLPWSECPVVRNGSVAVVEAECEKSSATTYFWYREALDISDSISESGGLNWKMTLCL 231

Qy     232 LV 233
           ||
Db     232 LV 233


RESULT 3
ABB06153
ID   ABB06153 standard; Protein; 266 AA.
XX
```

AC ABB06153;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS protein sequence SEQ ID NO:245.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR N-PSDB; ABL39807.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 6; Page 280-281; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 266 AA;

Query Match 7.7%; Score 56; DB 23; Length 266;
Best Local Similarity 100.0%; Pred. No. 4.9e-49;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELD 56
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 43 MPKNSKVTQREHSSEHVTESVADLLALEEPVDYKQSVLNVAGEAGGKQKAVEEELD 98

RESULT 4
AAO05876
ID AAO05876 standard; Protein; 122 AA.
XX
AC AAO05876;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 19768.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI85807.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 19768; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 122 AA;

Query Match 7.2%; Score 52; DB 22; Length 122;
Best Local Similarity 100.0%; Pred. No. 3.3e-45;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 354 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEM 405
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 20 VLGFKANIMNEKCVVENAEKILGYLNTNVLSRDLIPPHVNFSHLTTKDYMEM 71

RESULT 5
ABB11188
ID ABB11188 standard; peptide; 143 AA.
XX

AC ABB11188;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human neurotransmitter transporter homologue, SEQ ID NO:1558.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA08432.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 20; Page 154; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 143 AA;

Query Match 6.1%; Score 44; DB 22; Length 143;
Best Local Similarity 100.0%; Pred. No. 7.4e-37;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 505 FLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYG 548
||||||||||||||||||||||||||||||||||||||||||||
Db 71 FLVGLLFVQRSGNYFVTMFDDYSATLPLTLIVILENIAVAWIYG 114

RESULT 6
AAM24377
ID AAM24377 standard; Protein; 127 AA.
XX
AC AAM24377;
XX
DT 12-OCT-2001 (first entry)
XX
DE Rat EST encoded protein SEQ ID NO: 1902.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition.
XX
OS Rattus sp.
XX
PN WO200154477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02687.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.
DR N-PSDB; AAH99036.
XX
PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
PS Claim 20; Page 1225; 1275pp; English.
XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.
XX
SQ Sequence 127 AA;

Query Match 5.1%; Score 37; DB 22; Length 127;

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e-29;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     253 MYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKL 289
           |||||||||||||||||||||||||||||||||||||
Db       1 MYFSSLFPYVVLACFLVRGLLLRGAVDGILHMFTPKL 37


RESULT 7
AAM24378
ID   AAM24378 standard; Protein; 83 AA.
XX
AC   AAM24378;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Rat EST encoded protein SEQ ID NO: 1903.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition.
XX
OS   Rattus sp.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   N-PSDB; AAH99037.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 20; Page 1225; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a
CC   protein of the invention.
XX
SQ   Sequence   83 AA;

  Query Match             3.6%;  Score 26;  DB 22;  Length 83;
  Best Local Similarity   100.0%;  Pred. No. 2e-18;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     571 MWKFVSPLCMAVLTTASIIQLGVTPP 596
           ||||||||||||||||||||||||||
Db       1 MWKFVSPLCMAVLTTASIIQLGVTPP 26


RESULT 8
ABB11163
ID   ABB11163 standard; peptide; 185 AA.
XX
AC   ABB11163;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human orphan transporter homologue, SEQ ID NO:1533.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03800.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   N-PSDB; ABA08407.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 20; Page 151-152; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
```

CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 185 AA;

Query Match 3.0%; Score 22; DB 22; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 511 FVQRSGNYFVTMFDDYSATLPL 532
||||||||||||||||||||||
Db 30 FVQRSGNYFVTMFDDYSATLPL 51

RESULT 9
ABB57432
ID ABB57432 standard; Protein; 188 AA.
XX
AC ABB57432;
XX
DT 15-MAR-2002 (first entry)
XX
DE Human secreted protein encoding polypeptide SEQ ID NO 78.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200183510-A1.
XX
PD 08-NOV-2001.
XX
PF 26-APR-2001; 2001WO-US13318.
XX
PR 02-MAY-2000; 2000US-201194P.
PR 16-JUN-2000; 2000US-212142P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis G, Ruben SM, Rosen CA;
XX
DR WPI; 2002-121886/16.
XX
PT An isolated nucleic acid molecule encoding a human secreted protein
PT (SP) is useful in preventing, treating or ameliorating a disorder e.g.,
PT Alzheimer's disease and cancers -
XX
PS Disclosure; Page 473; 496pp; English.
XX
CC The invention relates to novel genes (ABL01566-ABL01594) and proteins
CC (ABB57394-ABB57456) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
SQ Sequence 188 AA;

Query Match 3.0%; Score 22; DB 23; Length 188;
Best Local Similarity 100.0%; Pred. No. 6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 511 FVQRSGNYFVTMFDDYSATLPL 532
||||||||||||||||||||||
Db 142 FVQRSGNYFVTMFDDYSATLPL 163

RESULT 10
ABB57433
ID ABB57433 standard; Protein; 188 AA.
XX
AC ABB57433;
XX
DT 15-MAR-2002 (first entry)
XX
DE Human secreted protein encoding polypeptide SEQ ID NO 79.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200183510-A1.
XX
PD 08-NOV-2001.
XX
PF 26-APR-2001; 2001WO-US13318.
XX
PR 02-MAY-2000; 2000US-201194P.
PR 16-JUN-2000; 2000US-212142P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis G, Ruben SM, Rosen CA;
XX
DR WPI; 2002-121886/16.
XX
PT An isolated nucleic acid molecule encoding a human secreted protein
PT (SP) is useful in preventing, treating or ameliorating a disorder e.g.,
PT Alzheimer's disease and cancers -
XX
PS Disclosure; Page 473-474; 496pp; English.
XX
CC The invention relates to novel genes (ABL01566-ABL01594) and proteins
CC (ABB57394-ABB57456) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
SQ Sequence 188 AA;

```
  Query Match             3.0%;  Score 22;  DB 23;  Length 188;
  Best Local Similarity   100.0%;  Pred. No. 6e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     142 FVQRSGNYFVTMFDDYSATLPL 163


RESULT 11
AAO21214
ID   AAO21214 standard; Protein; 259 AA.
XX
AC   AAO21214;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Sodium-dependent neurotransmitter transporter protein.
XX
KW   Antiparkinsonian; nootropic; neuroprotective; tranquiliser; vulnerary;
KW   cerebroprotective; human sodium-dependent neurotransmitter transporter;
KW   hybridising; central; peripheral nervous system disease; brain injury;
KW   cerebrovascular disease; Parkinson's disease; corticobasal degeneration;
KW   motor neuron disease; dementia; multiple sclerosis; post-stroke;
KW   traumatic brain injury; stroke; post-traumatic brain injury;
KW   small-vessel cerebrovascular disease; Alzheimer's disease; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200229048-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-EP11440.
XX
PR   05-OCT-2000; 2000US-237689P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Kohler RH;
XX
DR   WPI; 2002-426112/45.
DR   N-PSDB; AAL37763.
XX
PT   New human sodium-dependent neurotransmitter transporter polypeptide,
PT   the regulation of which is useful for treating central or peripheral
PT   nervous system diseases e.g. brain injuries, Parkinson's disease, and
PT   dementia  -
XX
PS   Claim 1; Fig 2; 98pp; English.
XX
CC   The invention relates to a purified human sodium-dependent
CC   neurotransmitter transporter polypeptide comprising a sequence of 259
CC   amino acids, given in the specification. A nucleic acid encoding the
CC   protein is useful for detecting a polynucleotide encoding the protein in
CC   a biological sample by hybridising the protein-encoding polynucleotide to
CC   a nucleic acid material of a biological sample to form a hybridisation
CC   complex, and detecting the hybridisation complex. The protein of the
CC   invention and its encoding gene are useful for screening for agents which
CC   decrease the activity of sodium-dependent neurotransmitter transporter
CC   polypeptide by contacting the test compound with the protein or its gene
CC   and detecting binding of the test compound to the protein or its gene. A
CC   vector comprising the polynucleotide of the invention or a reagent that
CC   modulates the activity of the protein or its gene is useful for the
CC   preparation of a medicament for modulating the activity of sodium-
CC   dependent neurotransmitter transporter in a disease. The reagent is
CC   useful for treating diseases such as central or peripheral nervous system
CC   disease, where the central nervous system disorders are selected from
CC   brain injuries, cerebrovascular diseases and their consequences,
CC   Parkinson's disease, corticobasal degeneration, motor neuron disease,
CC   dementia, including multiple sclerosis, traumatic brain injury, stroke,
CC   post-stroke, post-traumatic brain injury, small-vessel cerebrovascular
CC   disease, and Alzheimer's disease. The polynucleotide of the invention can
CC   be used in gene therapy. This sequence represents the 259 amino acid
CC   sodium-dependent neurotransmitter transporter protein of the invention.
XX
SQ   Sequence   259 AA;

  Query Match             3.0%;  Score 22;  DB 23;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 8e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     220 FVQRSGNYFVTMFDDYSATLPL 241


RESULT 12
AAG64743
ID   AAG64743 standard; Protein; 392 AA.
XX
AC   AAG64743;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human sodium neurotransmitter cotransporter protein 8,43.
XX
KW   Human; sodium neurotransmitter cotransporter protein 8,43; cytostatic;
KW   virucidal; immunomodulatory; antiinflammatory; haemostatic;
KW   malignant tumour; haemopathy; HIV infection; immunological disease;
KW   inflammatory condition.
XX
OS   Homo sapiens.
XX
PN   WO200153496-A1.
XX
PD   26-JUL-2001.
XX
PF   15-JAN-2001; 2001WO-CN00020.
XX
PR   21-JAN-2000; 2000CN-0111469.
XX
PA   (BIOD-) BIODOOR GENE TECHNOLOGY LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-465376/50.
DR   N-PSDB; AAH73714.
XX
PT   Human sodium neurotransmitter cotransporter protein 8,43 and encoded
PT   polynucleotide, applicable in diagnosis and treatment of malignant
PT   tumor, hemopathy, HIV infection, immunological diseases and various
PT   inflammations  -
XX
PS   Claim 1; Page 33-34; 40pp; Chinese.
XX
CC   This invention relates to human sodium neurotransmitter cotransporter
CC   protein 8,43 and the cDNA encoding it. Included in the invention is a
CC   vector containing the cDNA, a host cell transformed with the vector and
CC   an antibody targeting the protein. The protein, polynucleotide, antibody,
CC   or ant/agonist can be used to achieve cytostatic; virucidal;
CC   immunomodulatory; antiinflammatory or haemostatic activity. They can be
CC   used in the diagnosis and treatment of malignant tumours, haemopathy, HIV
CC   infection, immunological diseases and various inflammatory conditions.
CC   The present sequence represents the human sodium neurotransmitter
CC   cotransporter protein 8,43.
XX
SQ   Sequence   392 AA;

  Query Match             3.0%;  Score 22;  DB 22;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
```

```
Db    220 FVQRSGNYFVTMFDDYSATLPL 241


RESULT 13
AAO21215
ID   AAO21215 standard; Protein; 729 AA.
XX
AC   AAO21215;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Protein sequence identified by SwissProt Accession No: Q08469.
XX
KW   Antiparkinsonian; nootropic; neuroprotective; tranquiliser; vulnerary;
KW   cerebroprotective; human sodium-dependent neurotransmitter transporter;
KW   hybridising; central; peripheral nervous system disease; brain injury;
KW   cerebrovascular disease; Parkinson's disease; corticobasal degeneration;
KW   motor neuron disease; dementia; multiple sclerosis; post-stroke;
KW   traumatic brain injury; stroke; post-traumatic brain injury;
KW   small-vessel cerebrovascular disease; Alzheimer's disease; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200229048-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-EP11440.
XX
PR   05-OCT-2000; 2000US-237689P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Kohler RH;
XX
DR   WPI; 2002-426112/45.
XX
PT   New human sodium-dependent neurotransmitter transporter polypeptide,
PT   the regulation of which is useful for treating central or peripheral
PT   nervous system diseases e.g. brain injuries, Parkinson's disease, and
PT   dementia -
XX
PS   Disclosure; Fig 3; 98pp; English.
XX
CC   The invention relates to a purified human sodium-dependent
CC   neurotransmitter transporter polypeptide comprising a sequence of 259
CC   amino acids, given in the specification. A nucleic acid encoding the
CC   protein is useful for detecting a polynucleotide encoding the protein in
CC   a biological sample by hybridising the protein-encoding polynucleotide to
CC   a nucleic acid material of a biological sample to form a hybridisation
CC   complex, and detecting the hybridisation complex. The protein of the
CC   invention and its encoding gene are useful for screening for agents which
CC   decrease the activity of sodium-dependent neurotransmitter transporter
CC   polypeptide by contacting the test compound with the protein or its gene
CC   and detecting binding of the test compound to the protein or its gene. A
CC   vector comprising the polynucleotide of the invention or a reagent that
CC   modulates the activity of the protein or its gene is useful for the
CC   preparation of a medicament for modulating the activity of sodium-
CC   dependent neurotransmitter transporter in a disease. The reagent is
CC   useful for treating diseases such as central or peripheral nervous system
CC   disease, where the central nervous system disorders are selected from
CC   brain injuries, cerebrovascular diseases and their consequences,
CC   Parkinson's disease, corticobasal degeneration, motor neuron disease,
CC   dementia, including multiple sclerosis, traumatic brain injury, stroke,
CC   post-stroke, post-traumatic brain injury, small-vessel cerebrovascular
CC   disease, and Alzheimer's disease. The polynucleotide of the invention can
CC   be used in gene therapy. This sequence represents the protein sequence
CC   identified by SwissProt Accession No: Q08469 relating to the sodium-
CC   dependent neurotransmitter transporter protein of the invention.
XX
SQ   Sequence   729 AA;

  Query Match             3.0%;  Score 22;  DB 23;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     511 FVQRSGNYFVTMFDDYSATLPL 532
           ||||||||||||||||||||||
Db     512 FVQRSGNYFVTMFDDYSATLPL 533


RESULT 14
AAG67159
ID   AAG67159 standard; Protein; 730 AA.
XX
AC   AAG67159;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human 579 transporter polypeptide.
XX
KW   Human; transporter; 20685; 579; 17114; 23821; 33894; 32613;
KW   vesicular monoamine transporter; neurotransmitter-symporter;
KW   ABC transporter; sulfate transporter; neurological disorder;
KW   central nervous system disorder; Parkinson's disease; depression; pain;
KW   infectious disease; cell proliferative disorder; cancer; blood disorder;
KW   immune disorder; inflammatory disorder; spleen disorder; lung disorder;
KW   Hodgkin's disease; Niemann-Pick disease; chronic bronchitis; ischemia;
KW   colon disorder; cirrhosis; uterus disorder; endometrium disorder;
KW   endometrial stromal tumour; brain disorder; T-cell disorder; anemia;
KW   Sjogren syndrome; skin disorder; lupus erythematosus; heart disorder;
KW   haematopoietic stem cell; Alzheimer's disease; myocardial infarction;
KW   blood vessel; Kawasaki syndrome; red cell disorder; thymus disorder;
KW   B-cell disorder; kidney disorder; glomerulonephritis; breast disorder;
KW   testis disorder; thyroid disorder; Graves disease; pancreatitis;
KW   skeletal muscle disorder; tumour; pancreas disorder;
KW   small intestine disorder; celiac sprue.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          61..659
FT                   /note= "sodium neurotransmitter-symporter domain"
FT   Domain          70..87
FT                   /note= "transmembrane domain"
FT   Domain          98..117
FT                   /note= "transmembrane domain"
FT   Domain          140..164
FT                   /note= "transmembrane domain"
FT   Domain          228..244
FT                   /note= "transmembrane domain"
FT   Domain          253..275
FT                   /note= "transmembrane domain"
FT   Domain          306..323
FT                   /note= "transmembrane domain"
FT   Domain          334..358
FT                   /note= "transmembrane domain"
FT   Domain          458..479
FT                   /note= "transmembrane domain"
FT   Domain          496..513
FT                   /note= "transmembrane domain"
FT   Domain          527..550
FT                   /note= "transmembrane domain"
FT   Domain          575..594
FT                   /note= "transmembrane domain"
FT   Domain          617..639
FT                   /note= "transmembrane domain"
XX
PN   WO200164875-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06374.
XX
PR   29-FEB-2000; 2000US-0185906.
XX
```

PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Glucksmann MA;
XX
DR WPI; 2001-550178/61.
DR N-PSDB; AAH75186.
XX
PT Novel human transporter polypeptides useful for treating and diagnosing
PT Parkinson's disease, Hodgkin disease, glomerulonephritis, myocardial
PT infarction, Grave's disease, Alzheimer's disease, anemia, asthma and
PT tumours -
XX
PS Claim 9; Fig 9A-C; 259pp; English.
XX
CC The present sequence represents a human transporter polypeptide. The
CC specification describes 20685, 579, 17114, 23821, 33894 or 32613 human
CC transporter polypeptides. The 20685 transporter is similar to vesicular
CC monoamine transporters. The 579 transporter is similar to
CC neurotransmitter-symporters. The 17114 transporter is similar to ABC
CC transporters. The 32613 transporter is similar to sulfate transporters.
CC The transporter polypeptides and polynucleotides are useful for treating
CC and diagnosing neurological and central nervous system disorders (e.g.
CC Parkinson's disease, depression, pain), infectious disease, cell
CC proliferative disorders (e.g., cancer), blood disorders, and immune and
CC inflammatory disorders. They are also useful for treating and
CC diagnosing disorders involving the spleen (e.g., Hodgkin disease,
CC Niemann-Pick disease), lung (e.g., chronic bronchitis), colon
CC (cirrhosis), uterus and endometrium (e.g., endometrial stromal tumours),
CC brain (e.g., ischemia), T-cells (e.g., Sjogren syndrome), skin (lupus
CC erythematosus), haematopoietic stem cells (e.g, Alzheimer's disease),
CC heart (e.g., myocardial infarction), blood vessels (e.g., Kawasaki
CC syndrome), red cells (e.g., anemias), disorders involving thymus,
CC B-cells, kidney (e.g., glomerulonephritis), disorders involving breast,
CC testis, epididymis, prostate, thyroid (e.g., Graves disease), disorders
CC involving skeletal muscle (e.g, tumour), pancreas (e.g., pancreatitis),
CC small intestine (e.g., celiac sprue), disorders related to reduced
CC platelet number and ovary.
XX
SQ Sequence 730 AA;

Query Match 3.0%; Score 22; DB 22; Length 730;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 511 FVQRSGNYFVTMFDDYSATLPL 532
||||||||||||||||||||||
Db 512 FVQRSGNYFVTMFDDYSATLPL 533

RESULT 15
AAY72908
ID AAY72908 standard; Protein; 730 AA.
XX
AC AAY72908;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human NTT7 protein.
XX
KW Human; NTT7 protein; therapy; psychiatric disorder; anxiety; depression;
KW schizophrenia; phobia; panic disorder; obsessive compulsive disorder;
KW Parkinson's disease; central nervous system disorder; cerebroprotective;
KW neurological disorder; stroke; pain; neuropathic pain; sleep disorder;
KW tranquilliser; analgesic; neuroleptic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200119854-A2.
XX
PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-GB03541.
XX
PR 15-SEP-1999; 99GB-0021833.
PR 17-MAR-2000; 2000GB-0006545.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Hill J, Duckworth DM, Farmer M, Pangalos M;
XX
DR WPI; 2001-244780/25.
DR N-PSDB; AAD03039.
XX
PT Novel NTT7 polypeptide useful for treating anxiety, depression,
PT schizophrenia, phobia, Parkinson's disease, stroke, pain, and
PT psychiatric, panic, neurological, central nervous system, obsessive
PT compulsive and sleep disorders -
XX
PS Claim 3; Page 20; 27pp; English.
XX
CC The present sequence is human NTT7 protein encoded by a cDNA. NTT7 is
CC thought to be a member of the neurotransmitter family of polypeptides.
CC NTT7 sequences are useful for treating psychiatric disorders, anxiety,
CC depression, schizophrenia, phobias, panic disorder, obsessive compulsive
CC disorder, Parkinson's disease, central nervous system disorders, stroke,
CC neurological disorders, pain, neuropathic pain, sleep disorders, and
CC diseases in which neurotransmitters are implicated. NTT7 sequences are
CC useful for screening antagonists and agonists of NTT7 and as vaccines
CC for inducing immunological response in a mammal. NTT7 sequences and its
CC antibodies are useful to configure screening methods for detecting the
CC effect of added compounds on the production of mRNA and polypeptide in
CC cells. NTT7 is useful in conventional low capacity screening methods and
CC also in high-throughput screening (HTS) formats and is useful for
CC identifying membrane bound or soluble receptors. NTT7 antibody is useful
CC to isolate or identify clones expressing NTT7 or to purify NTT7 by
CC affinity chromatography. NTT7 is useful as diagnostic reagent for
CC detecting mutations in the associated gene, and for chromosome
CC localisation studies and tissue expression studies.
XX
SQ Sequence 730 AA;

Query Match 3.0%; Score 22; DB 22; Length 730;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 511 FVQRSGNYFVTMFDDYSATLPL 532
||||||||||||||||||||||
Db 512 FVQRSGNYFVTMFDDYSATLPL 533

Search completed: April 21, 2003, 14:09:56
Job time : 43 secs